Fall 12-23-2015

A Genetic and Biochemical Analysis of *Sulfolobus* Spindle-Shaped Virus 1

Eric Alexander Iverson
Portland State University

Recommended Citation
Iverson, Eric Alexander, "A Genetic and Biochemical Analysis of *Sulfolobus* Spindle-Shaped Virus 1" (2015). *Dissertations and Theses.* Paper 2641.
https://doi.org/10.15760/etd.2637

A Genetic and Biochemical Analysis of *Sulfolobus* Spindle-Shaped Virus 1

by

Eric Alexander Iverson

A dissertation submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy
in
Biology

Dissertation Committee:
Kenneth Stedman, Chair
Michael Bartlett
Justin Courcelle
John J. Perona
Jeffrey Singer

Portland State University
2015

Abstract

Viruses infecting the Archaea exhibit a tremendous amount of morphological and genetic diversity. This is especially true for crenarchaeal viruses from the family *Fuselloviridae*, which possess spindle-shaped capsids and genomes that harbor a great number of uncharacterized genes. The functions of these unidentified gene products are of interest as they have the potential to provide valuable insights into the fusellovirus infection cycle and archaeal viruses in general. In an effort to better characterize the genetic requirements of the *Fuselloviridae*, we have performed genetic and biochemical experiments using the best studied fusellovirus, *Sulfolobus* spindle-shaped virus 1 (SSV1).

A comprehensive genetic analysis of SSV1 was conducted using long inverse PCR and transposon mutagenesis. The results of this work illustrate that SSV1 is highly tolerant of mutagenesis. A robust protocol for the purification of recombinant VP2 protein from *E. coli* was developed and should be useful for future studies aimed at characterizing the biochemical and structural characteristics of this SSV1 structural protein. Finally, the first insights into a fusellovirus infection are presented and provide the framework for a detailed characterization of the fusellovirus infection cycle. The results and significance of this work are presented in the chapters that follow.

Acknowledgements

I'd like to thank the members of my committee: Dr. Michael Bartlett, Dr. Justin Courcelle, Dr. John Perona, and Dr. Jeffrey Singer. Their constructive criticism and insights were immensely helpful during this process.

I'd especially like to thank my PhD advisor Dr. Ken Stedman for his contributions that are much too numerous to list here. Thank you for giving me the opportunity to work in your lab, for staying patient with me, and teaching me so many things over the past few years. I truly loved working with you, it has been a blast.

Thank you to all of the Stedman lab "cronies" throughout the years: Dr. Geoff Diemer (your advice during my first two years was monumental), Dr. Jim Laidler, Mikaela Selby, Dr. Jennifer Kyle, George Kasun, and many, many more. In particular, I'd like to thank David Goodman whose brilliant attitude, sense of humor, and scientific contributions were a humongous help throughout.

Last and certainly not least I'd like to thank my wife, Dr. Katherine Iverson. Thank you for humoring me with all of my virus talk over the past 6 years and for understanding that "I'll be done in 30 minutes" actually means "I'll *probably* be done in 3 hours". I absolutely could not have done this without your constant support. Thank you and I love you very much.

This research was funded by:

Portland State University and grants from the National Science Foundation, Grants MCB: 0702020 and 1243963.

NSF Eastern Asia and Pacific Summer Institute (EAPSI) Fellowship, Award number: IIA-1414642

TABLE OF CONTENTS

ABSTRACT i

ACKNOWLEDGEMENTS ii

LIST OF TABLES viii

LIST OF FIGURES ix

CHAPTER 1: INTRODUCTION 1

THE ARCHAEA AND THEIR VIRUSES 1
ARCHITECTURE OF ARCHAEAL VIRIONS: A MIX OF OLD AND NEW 4
SULFOLOBUS SPINDLE-SHAPED VIRUS 1 (SSV1) 5
COMPARATIVE GENOMICS OF ARCHAEAL VIRUSES 11
COMPARATIVE GENOMICS OF FUSELLOVIRUSES 13
ARCHAEAL VIRUS PROTEIN STRUCTURE AND FUNCTION 16
GENE EXPRESSION IN VIRUSES OF *SULFOLOBUS* 20
DNA REPLICATION IN HYPERTHERMOPHILIC ARCHAEA AND VIRUSES 25
CRISPR-CAS SYSTEMS: OVERVIEW AND INTERACTIONS WITH ARCHAEAL VIRUSES 27
CONCLUSION 32

CHAPTER 2: A GENETIC STUDY OF SSV1, THE PROTOTYPICAL FUSELLOVIRUS 34

ABSTRACT 34
INTRODUCTION 35
MATERIALS AND METHODS 39
Culture conditions 39
Purification of DNA 39
Construction of SSV1 deletion mutants 40
Electroporation of Sulfolobus 41
Screen for functional infectious virus/halo assay 41
Growth curves 42
Transmission electron microscopy (TEM) 42
RESULTS 43
SSV1 is infectious without the vp2 *gene* 43
SSV1 constructs lacking the conserved ORF b129 *do not appear to make infectious viruses* 45
SSV1 lacking ORF d244 *is infectious but has a novel phenotype* 47
DISCUSSION 49
The putative DNA packaging protein VP2 is not required for SSV1 function 49
The product of ORF b129 *appears to be essential for SSV1 infectivity* 50
*Transfection with SSV-Δ*d244 *produces virus and retards cell growth* 51
SUMMARY AND OUTLOOK 53

CHAPTER 3: A COMPREHENSIVE GENETIC SURVEY OF THE FUSELLOVIRUS SSV1 REVEALS A MALLEABLE GENOME THAT TOLERATES DELETION OF THE MINOR CAPSID GENE..... 54

ABSTRACT 54
INTRODUCTION 55
MATERIALS AND METHODS 62
Purification of SSV1 DNA from Sulfolobus 62
Purification of LIPCR template DNA from E. coli 64
Long inverse PCR (LIPCR) 64
Transposon mutagenesis 65
Preparation of chemically competent E. coli *for transformation of SSV1 deletion mutants* 66
Isolation and identification of transposon and deletion mutants 67
Complementation in cis of SSV1 mutants 68
Transformation of Sulfolobus 68
Halo Assay to check for infectious virus production 69
Confirmation of infectious SSV DNA 70
Transmission electron microscopy (TEM) 70
RESULTS 72
Construction of specific SSV1 mutants by LIPCR 72
Construction of SSV1 EZTn5 transposon insertion mutants 80
Summary of SSV1 mutagenesis 80
The moderately conserved ORFs of the T6 transcript 82
Both the N- and C-termini of b129 *are essential for SSV1 infectivity* 84
The monocistronic transcripts T3 and Tx 84
The fusellovirus core is intolerant of mutagenesis 85
The T5 transcript region is almost entirely dispensable 86
Integrase and e54 *deletion mutants exhibit a variable host range* 87
ORFS b49 *and* f55 *and the predicted origin of replication* 88
The minor capsid gene vp3 *is highly conserved but non-essential* 89
Homologous recombination can result in deletion of the vp3 gene 91
Mutagenesis of the major capsid gene vp1 93
Complementation in cis of SSV1 deletion mutants 95
DISCUSSION 96
Overview of SSV1 mutagenesis 96
The T5 transcript is almost entirely dispensable 97
The moderately conserved ORFs of the T6 transcript 99
The SSV1 integrase *gene may be essential in some hosts* 100
The Tx and T3 transcripts 102
The fusellovirus core is intolerant of mutagenesis 107
ORFs b49 *and* f55 *and the predicted origin of replication* 108
The vp3 *minor capsid gene is not essential* 111
Mutagenesis of the major capsid gene vp1 113

CONCLUSIONS 115

CHAPTER 4: PURIFICATION AND CHARACTERIZATION OF THE DNA BINDING PROTEIN VP2 FROM SSV1 118

ABSTRACT 118
INTRODUCTION 118
MATERIALS AND METHODS 121
Cloning the vp2 *gene of SSV1 into pET30 expression vector* 121
Removal of N-terminal tag region from pET30_VP2 expression vector 122
Overexpression of 6-His VP2 protein 123
Initial purification of 6-His VP2 from E. coli 124
Ni-NTA chromatography 125
Heparin affinity chromatography 125
Size exclusion chromatography 126
Circular dichroism (CD) spectroscopy 126
Electrophoretic mobility shift assays (EMSAs) 127
Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) 127
Bradford Assay 128
Trichloroacetic acid (TCA) precipitation 128
RESULTS 128
Cloning of SSV1 vp2 *and expression of recombinant 6-His VP2* 128
Ni-NTA chromatography of recombinant 6-His VP2 130
Heparin affinity chromatography 134
SEC purification indicates 6-His VP2 exists as a monomer in solution 136
6-His VP2 is folded following purification 137
DNA binding by 6-His VP2 138
DISCUSSION AND CONCLUSIONS 139
Purification of recombinant 6-His VP2 140
6-His VP2 exists as a monomer in solution 141
6-His VP2 binds non-specifically to linear and circular dsDNA 142
Comparison of VP2 homologues 143
Putative role of VP2 in the Fuselloviridae 146

CHAPTER 5: FIRST INSIGHTS INTO THE KINETICS OF THE FUSELLOVIRUS REPLICATION CYCLE 150

ABSTRACT 150
INTRODUCTION 150
MATERIALS AND METHODS 156
Growth of Sulfolobus *cultures* 156
Comparison of filtered and non-filtered virus supernatants 156
Initial purification of virus for one-step growth curves 157
SSV1 production in stably infected cells 157
Low MOI infections and one-step growth curve experiments 158
Calculation of virus burst size 160

Experiments with "spent" growth medium 161
*UV-inactivation of SSV1(*e178*::Topo) virions* 161
Plaque assay 163
RESULTS 164
Filtration of virus results in a 100-fold reduction in infectivity 164
Newly transfected Sulfolobus *yields low viral titers* 165
Stably infected cultures produce reasonably high (10^7 – 10^8 PFU/mL) titers 166
*Infections with SSV1 and SSV1(*e178*::Topo) at low MOI* 168
Infection with SSV9 at low MOI 171
One-step growth curves at a MOI of 5 173
*One-step growth curves of SSV1, SSV1(*e178*::Topo), and the SSV1 mutant SSV1Δ*d244*(*e178*::Topo)* 174
The effects of spent media and UV-inactivated virus on S441 growth 177
DISCUSSION 180
Filtration of SSV1 results in a decrease in infectivity 180
Newly transfected Sulfolobus *results in the production of low viral titers* 181
Stably infected cultures produce moderately high (10^7 – 10^8 PFU/mL) titers 181
Infection of Sulfolobus *at a low MOI* 182
One-step growth curves in Sulfolobus *at a MOI of 5* 184
Comparisons of one-step growth curves 184
Virus burst size and latent period 185
*Cell growth during SSV1 and SSV1(*e178*::Topo) one-step growth curves* 187
*SSV1Δ*d244*(*e178*::Topo)-infected cells display a distinct phenotype* 189
High initial starting titers in one-step growth curves 191
Appearance of cell debris in high MOI infections 191
Infection with SSV9 at low MOI 192
Comparisons to other archaeal viruses 197
CONCLUSIONS 198

CHAPTER 6: SUMMARY AND FUTURE DIRECTIONS 201

THE ROLE OF VP2 IN THE *FUSELLOVIRIDAE* 201
THE GENETIC CHARACTERIZATION OF SSV1 203
THE KINETICS OF A FUSELLOVIRUS INFECTION 208
CONCLUSION 210

REFERENCES CITED 212

APPENDIX A: SSV9 MUTAGENESIS 238

APPENDIX B: GENOME CONSERVATION IN THE *FUSELLOVIRIDAE* 246

List of Tables

Table 1-1. *List of fusellovirus genomes* 15
Table 1-2. *Open reading frames comprising the fusellovirus core* 16
Table 2-1. *Strains and plasmid vectors used in this work* 43
Table 2-2. *Oligonucleotides used in this work* 46
Table 3-1. Sulfolobus *and* E. coli *strains used in this work* 63
Table 3-2. *LIPCR primers used to construct SSV1 mutants* 71
Table 3-3. *List of plasmids and SSV1 shuttle vectors used in this work* 75
Table 3-4. *List of other primers used in this work* 79
Table 4-1. *Primers used for cloning the SSV1* vp2 *gene* 121
Table 4-2. *Buffers used in VP2 purification* 124
Table 4-3. *Properties of VP2 homologues and* S. solfataricus *P2 NAPs* 144
Table 5-1. *Strains, viruses, and plasmids used in this work* 158
Table 5-2. *Effect of filtration (0.45 μm) on SSV1 and SSV1(e178::Topo) supernatants* 165
Table 5-3. *Virus burst sizes* 177

List of Figures

Figure 1-1. *Transmission electron micrographs displaying morphological diversity of known archaeal viruses*.....5
Figure 1-2. *Transmission electron micrographs displaying rosette aggregates of SSV9 particles*.....10
Figure 1-3. *Map of the SSV1 genome*.....15
Figure 2-1. *Genome map of SSV1*.....37
Figure 2-2. *Typical growth inhibition of* S. solfataricus *on plates due to infectious virus*.....45
Figure 2-3. *Transmission electron micrographs of SSV particles*.....47
Figure 2-4. *Typical growth inhibition in liquid culture of virus constructs*.....48
Figure 3-1. *Genome map of SSV1*.....62
Figure 3-2. *Results of SSV1 mutagenesis*.....83
Figure 3-3. *B129 and F112 protein sequences*.....84
Figure 3-4. *The T5, T6, T_{ind}, and T_{lys} promoter region*.....89
Figure 3-5. *Electron micrographs of SSV1 structural mutants*.....90
Figure 3-6. *Analysis of functional insertion mutants EAI446, EAI452, and EAI476 within VP1/VP3 structural gene region*.....92
Figure 3-7. *Fusellovirus VP1 and VP3 proteins*.....95
Figure 3-8. *Alignment of putative SSV1-C124 homologues occupying a similar position in the fusellovirus genome*.....105
Figure 3-9. *Conservation of SSV1-A291 and the T3 transcript in the* Fuselloviridae.....106
Figure 4-1. *Amino acid sequences of recombinant VP2 proteins used in this work*.....129
Figure 4-2. *Overexpression of 6-His VP2 in BL21 DE3* E. coli**.**.....130
Figure 4-3. *Ni-NTA purification of 6-His VP2 under non-denaturing conditions*132
Figure 4-4. *Ni-NTA purification of 6-His VP2 using 7 M urea in the absence of NaCl results in incomplete removal of co-purifying nucleic acids*.....133
Figure 4-5. *Purification of 6-His VP2 by heparin column chromatography*.....135
Figure 4-6. *Size exclusion chromatography of 6-His VP2*.....136
Figure 4-7. *CD spectrum of 6-His VP2 at 25° C*.....137
Figure 4-8. *6-His VP2 binding of circular and EcoRI-digested pAJC97 (SSV1) DNA*.....139
Figure 4-9. *Multiple sequence alignments of VP2 homologues*.....146
Figure 5-1. *Growth and virus production of SSV1-infected S441 cells*.....167
Figure 5-2. *Growth curves of S441 cells infected with SSV1 and SSV1(*e178*::Topo) at low MOI*.....170
Figure 5-3. *Growth curve of S441 infected with SSV9 at a MOI of 0.1*.....172
Figure 5-4. *One-step growth curve of S441 infected with SSV1(*e178*::Topo) at a MOI of 5*.....174
Figure 5-5. *One-step growth curve of S441 cells infected with SSV1, SSV1(*e178*::Topo), and SSV1Δ*d244*(*e178*::Topo) at a MOI of 5*.....177

Figure 5-6. *The effect of spent growth medium and UV-inactivated SSV1(*e178*::Topo) on S441 cell growth* 179

Chapter 1: Introduction

The Archaea and their viruses

The pioneering work of Carl Woese in the late 1970's formally introduced the scientific community to the Archaea and triggered a paradigm shift in the way we looked at the interrelatedness of life on planet Earth (Woese and Fox, 1977; Albers et al., 2013). Later it was shown that the Archaea share traits with both Bacteria (compact genomes, genes arranged in operons, metabolism) and Eukarya (information processing systems) and as well as possess their own unique attributes (Zillig et al., 1991; Brochier-Armanet and Forterre 2011; Barry and Bell, 2006; Makarova and Koonin 2013). The Archaea are often thought to be exclusively extremophilic but it has become clear that in addition to populating some of the planet's most extreme environments, they can be found almost everywhere (Delong 1998; Schleper et al., 2005). However, it is their ability to thrive in the most extreme conditions on the planet that has made them very suitable subjects for studies into the origins of life and astrobiology. Furthermore, the similarities of their information processing systems with Eukaryotes and their general ease of manipulation has made them great surrogates in the laboratory (Cavicchioli 2011).

The advent of whole genome sequencing has greatly expanded our knowledge of microbial diversity, especially within the Archaea. The Archaea have been divided into three phyla (Crenarchaea, Euryarchaea, and Thaumarchaea) and three candidate phyla (Nanoarchaea, Korarchaea, and

Aigarchaeota) have been proposed (Brochier-Armanet and Forterre 2011). The Euryarchaea, Crenarchaea, and Nanoarchaea all include hyperthermophiles (Stetter et al., 2006). Members of the crenarchaeal genus *Sulfolobus* were some of the first hyperthermophiles to be isolated and are some of the best studied Archaea (Brock et al., 1972). *Sulfolobus* thrive in high temperature (70-80^{o}C) and low pH (2-4) hot-spring environments throughout the world (Whitaker et al., 2003). Because *Sulfolobus* grows aerobically it is much more straightforward to manipulate in the lab than most other hyperthermophiles, which are strict anaerobes. Furthermore, a number of diverse extrachromosomal genetic elements have been identified and isolated from many members of *Sulfolobus*, including some truly novel and unique viruses (Lipps et al., 2006; Prangishvili 2013; Wang et al., 2015).

Viruses infecting Archaea, particularly the Crenarchaea, harbor a tremendous amount of genomic and morphological diversity (Prangishvili et al., 2006; Krupovic et al., 2011). Archaeal virus genomes are linear or circular dsDNA, with the exception of two recently discovered ssDNA viruses, one of which encodes the largest ssDNA genome of any known virus (Mochizuki et al., 2012; Pietilä et al., 2009). No RNA virus infecting Archaea has been isolated to date but a metagenomic analysis of a hot spring in Yellowstone National Park has hinted at their possible existence (Bolduc et al., 2012). The vast majority of archaeal virus genes do not have recognizable homologues in public databases, even when compared to other archaeal viruses (Prangishvili et al., 2006).

Furthermore, a number of unique virion morphologies have been observed that do not appear in the other domains, including spindle-shaped, bottle-shaped, and droplet-shaped virions (Prangishvili et al., 2013). These fascinating and Archaea-specific architectures undoubtedly have a lot to teach about viral evolution as well as protein stability and protein folding in extreme environments (Snyder et al., 2015).

Virions with a spindle-shaped morphology have routinely been observed in geothermal and halophilic environments that are rich in Archaea (Luk et al., 2014; Prangishvili et al., 2013). These viruses have been divided into two viral families, the short-tailed *Fuselloviridae* (also known as SSVs) and the long-tailed *Bicaudaviridae* (ICTV 2014 release). An additional unassigned genus has been proposed, the *Salterprovirus*, and includes the halophilic virus His1 (Krupovic et al., 2014). Acidianus two-tailed virus (ATV), the only official member of the *Bicaudaviridae*, has been well studied and is best known for the unique development of two tails independently of and external to the host (Häring et al., 2005). ATV exhibits no significant genetic similarity with viruses in the *Fuselloviridae*, including the major capsid protein (MCP) (Krupovic et al., 2014; Appendix A). Several viruses with morphologies similar to fuselloviruses have been isolated but appear to share no significant genetic identity with the fuselloviruses other than two moderately similar hydrophobic patches within the MCP. Because of the weakly similar MCPs, it has been proposed that these new viruses be classified as fuselloviruses (Krupovic et al., 2014).

Architecture of archaeal virions: A mix of old and new

Archaeal viruses display a tremendous amount of morphological diversity. In addition to the classical bacteriophage architectures (e.g. head-tail, icosahedral and helical morphologies), archaeal viruses exhibit numerous bizarre structures that have never before been observed in viruses from the other domains (Prangishvili 2013; Figure 1-1). This seems to be especially true for archaeal viruses isolated from thermophilic environments. While this may be attributable to sampling bias, perhaps the prevalence of unique morphologies in these hostile environments reflects some fundamental adaptation to life in these inhospitable conditions. Regardless, archaeal virus morphologies and the structures of the proteins that comprise them have been instrumental in altering the way we look at viral relationships among the three domains.

Due to the limited number of solved structures and practically non-existent sequence homology, it is unclear just how many unique protein folds are utilized by the MCPs (major capsid proteins) of archaeal viruses. However, it is clear that at least a few archaeal viruses utilize some of the more ubiquitous protein folds to form their capsids. The MCP of the icosahedral virus STIV1 exhibits a double jelly-roll motif that has been observed in the bacteriophage PRD1, eukaryotic NCLDVs, and human adenovirus (Khayat et al., 2005; Veesler et al., 2013). More recently, a single jelly-roll motif was observed in the MCP of the haloarchaeal virus SH1 (Jäälinoja et al., 2008). Additionally, the HK97 fold found in the bacteriophage HK97 and eukaryotic herpesvirus was also recently

observed in the haloarchaeal virus HSTV1 (Pietilä et al., 2013). The presence of conserved virion architectures reveals potential hidden evolutionary relationships and implies the existence of a common ancestor prior to the split of the three domains (Khayat et al., 2005; Benson et al., 2004).

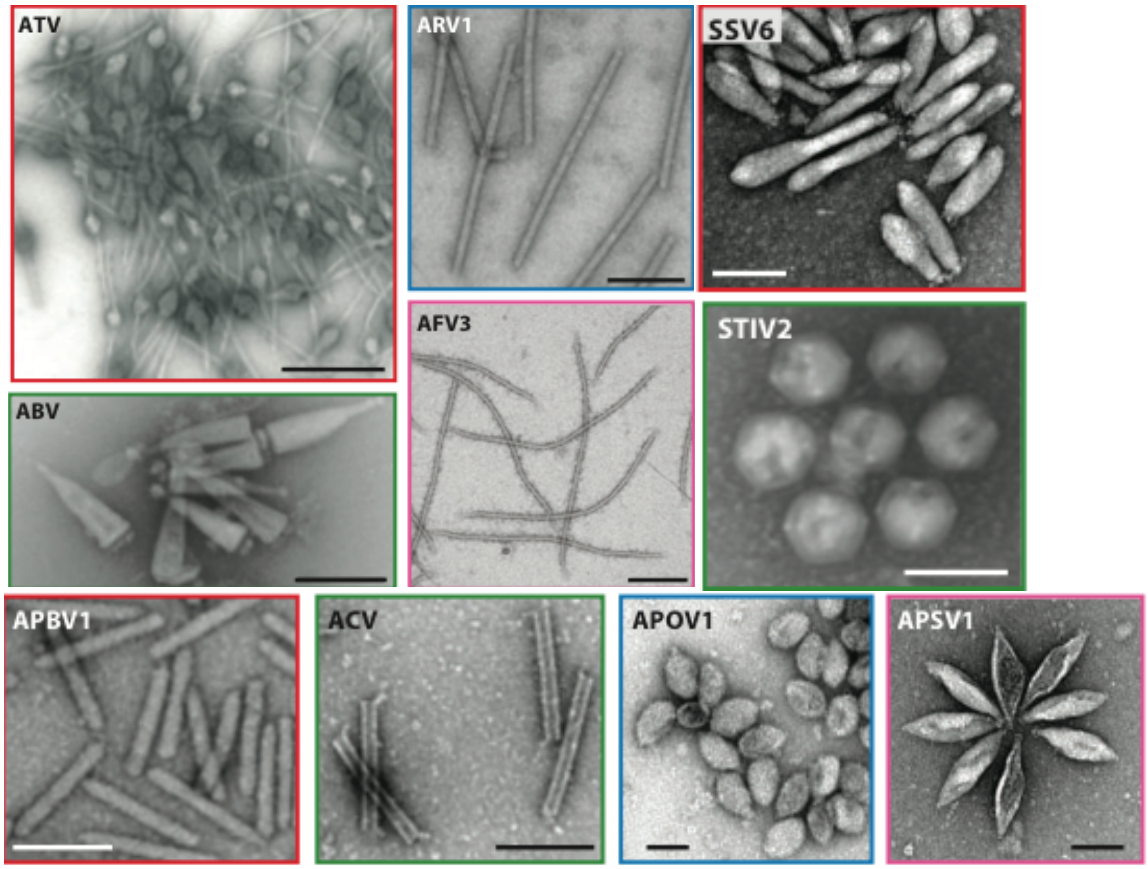


Figure 1-1: Transmission electron micrographs displaying morphological diversity of known archaeal viruses.
Acidianus two-tailed virus (ATV); *Acidianus* rod-shaped virus 1 (ARV1); *Acidianus* bottle-shaped virus (ABV); *Acidianus* filamentous virus 3 (AFV3); *Aeropyrum pernix* bacilliform virus 1 (APBV1); *Aeropyrum* coil-shaped virus (ACV); *Sulfolobus* spindle-shaped virus 6 (SSV6); *Sulfolobus* turreted icosahedral virus 2 (STIV2); *Aeropyrum pernix* ovoid virus 1 (APVO1); *Aeropyrum pernix* spindle-shaped virus 1 (APSV1). Bars, 100 nm. Image from Prangishvili 2013.

Sulfolobus spindle-shaped virus 1 (SSV1):

SSV1 (*Sulfolobus shibatae* virus 1, also *Sulfolobus* spindle-shaped virus 1) was the first-isolated and is probably the best-studied fusellovirus (Yeats et al.,

1982; Martin et al., 1984; Schleper et al., 1992; Contursi et al., 2014). The original host *S. shibatae* (isolated in Beppu, Japan and misidentified as *S. acidocaldarius*) was noticed to harbor double-stranded plasmid DNA (Yeats et al., 1982) and later found to produce lemon- or spindle-shaped particles following exposure to UV light (Martin et al., 1984). Schleper et al. later showed that *S. solfataricus* could be transfected with this extrachromosomal DNA, resulting in the integration of the DNA and the production of lemon-shaped particles (Schleper et al., 1992). Fuselloviruses have since been isolated from hot spring environments in Iceland (SSV2, SSV3, SSV4, SSV5, SSV6, and SSV7), Yellowstone National Park (SSV8), and Kamchatka, Russia (SSV9) (Prangishvili 2013). *Acidianus* spindle-shaped virus 1 (ASV1) was discovered in the process of sequencing the genome of *Acidianus brierleyi* and is the only identified fusellovirus shown to infect a non-*Sulfolobus* host (Redder et al., 2009). A fusellovirus genome was identified in a Mexican hot spring metagenome, however no virions were isolated and a few core fusellovirus genes appear to be missing (Servin-Garcidueñas et al., 2013; Table 1-2). A comprehensive analysis of fusellovirus host range determined that geography is not a reliable predictor of host susceptibility and found SSV1 to exhibit the narrowest host-range while SSV9 exhibits the broadest range (Ceballos et al., 2012; Held and Whitaker 2009).

The 15,465 bp SSV1 genome encodes 35 open reading frames (ORFs) and exists in a positively supercoiled topological state within the virion (Nadal et al.,

1986; Palm et al., 1992; Fusco et al., 2013; Figure 1-3). It is unclear how positive supercoiling arises or if other fusellovirus genomes are also positively supercoiled (Nadal et al., 1986). The SSV1 genome exists integrated into a host arginyl tRNA gene and also as a circular episome (Yeats et al., 1982). As discussed below, the function the vast majority of SSV1 proteins (and fusellovirus proteins in general) is unknown, although progress has been made using structural and comparative genomics (Lawrence et al., 2009; Prangishvili et al., 2006; Chapter 2). The SSV1 integrase is the only protein for which a definitive function can be assigned on the basis of sequence alone (Palm et al., 1991). Interestingly, SSV1 remained infectious when the *integrase* gene was deleted (Clore and Stedman 2006). However, SSV1-Δ*int* virus was rapidly outcompeted by the wild-type virus, suggesting that the universally conserved integrase plays an important role in the success of SSV1 (Clore and Stedman 2006).

Fusellovirus virions are ~100 nm x ~60 nm by negative stain transmission electron microscopy (TEM) and contain small tail fibers protruding from one end of the capsid that often bind with the tail fibers of other virions to form characteristic rosettes (Figure 1-2). A recent study illustrated that SSV1 aggregates were partially broken up by the addition of 1M NaCl and could be further disrupted by treatment with 1% ethanol, suggesting that inter-tail fiber interactions are ionic and hydrophobic in nature (Quemin et al., 2015). Interestingly, SSV6 and ASV1 virions display morphologically different tail fibers, resulting in virions that do not appear to form rosettes. Furthermore, SSV6 and

ASV1 exhibit pleiomorphic morphologies not observed in the other fuselloviruses (Redder et al., 2009).

The structure of SSV1 was solved to a resolution of 32 Å using cryo-EM microscopy and has given the first insights into the architecture and assembly of spindle-shaped viruses (Stedman et al., 2015). The capsid of SSV1 is composed of structural proteins VP1 and VP3 while the tail fiber is presumably formed by the VP4 protein (Reiter et al., 1987a; Menon et al., 2008; Redder et al., 2009; Quemin et al., 2015). The structure of the SSV1 virion indicates that the size of the SSV1 tail structure is of a similar size with pores in the glycoprotein S-layer of the host, strengthening the hypothesis that the tail fibers are used in host recognition and attachment (Stedman et al., 2015).

One additional protein was identified in SSV1 virions via mass spectrometry, the putative PD-(D/E)-XK-like nuclease D244, however, D244 was not identified in a more recent analysis (Menon et al., 2008; Menon et al., 2010; Quemin et al., 2015). Although not visible in the cryo-EM structure, host-derived tetraether lipids were identified in highly purified SSV1 virions, clearing up a long-standing debate on their presence in SSV1 virions (Quemin et al., 2015). VP1, VP3, and VP4 were all shown to be post-translationally glycosylated in purified SSV1 virions (Quemin et al., 2015). Although the site of glycosylation is unknown, each protein encodes two or more copies of the consensus sequence recognized by the host *Sulfolobus* glycosylation apparatus (Quemin et al., 2015). The major capsid protein of STIV1 is also known to be glycosylated, suggesting

glycosylation may be wide spread in the structural proteins of crenarchaeal viruses (Maaty et al., 2006).

A high contour level image of the SSV1 capsid allows for the approximate placement of seven potential capsomers in a hexagonal lattice. Interestingly, the distance between capsomers is only ~ 50 Å, significantly smaller than the distance between capsomers exhibiting more canonical jelly-roll or HK97 folds. This, along with the total absence of sequence homology from the SSV1 MCP, strongly suggests that the SSV1 MCP does not adopt either of these well known folds. Confirmation of this hypothesis awaits a higher resolution structure of SSV1 and/or the capsid proteins VP1 and VP3 (Stedman et al., 2015).

A novel viral release mechanism involving the construction of pyramid-like structures on the host cell membrane has been characterized for the two unrelated lytic crenarchaeal viruses, STIV1 and SIRV1 (Bize et al., 2009; Brumfield et al., 2009). The viral gene responsible for the production of these pyramid-like structures has been identified, however, no obvious homologue is present in the *Fuselloviridae* nor has any similar structure been reported or observed during a fusellovirus infection (personal observation). Host-derived tetraether lipids were recently found associated with SSV1 virions by mass spectrometry, strengthening the long-held hypothesis that SSV1 release occurs via budding (Quemin et al., 2015). However, the specifics of fusellovirus egress remain mostly unclear and await further research.

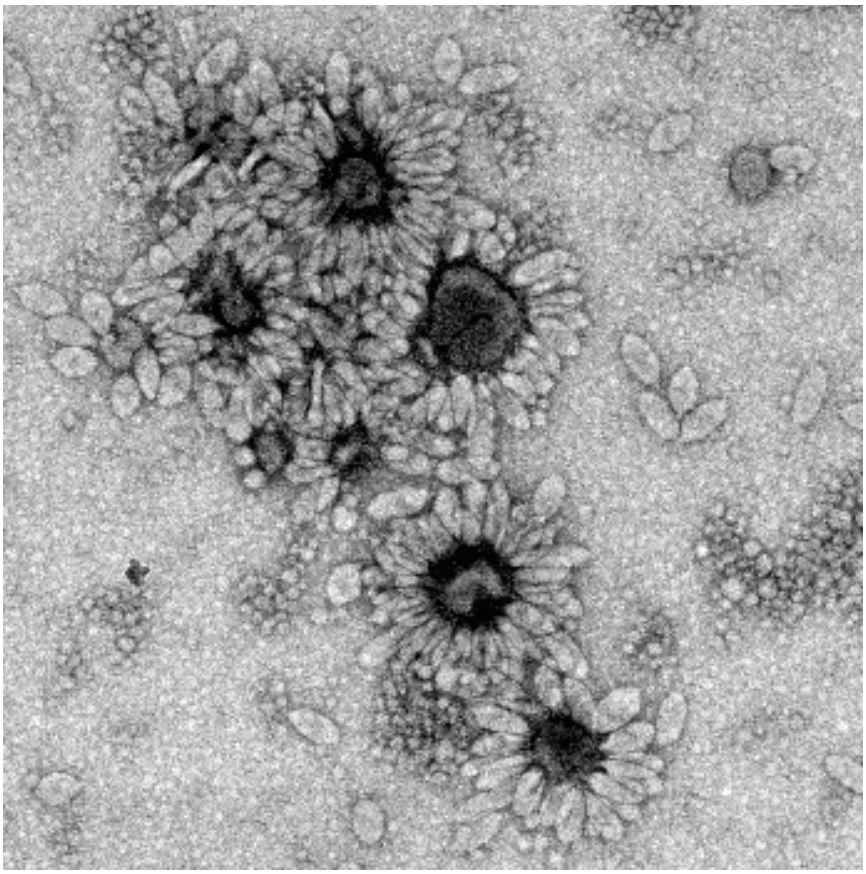


Figure 1-2: Transmission electron micrographs displaying rosette aggregates of SSV9 particles. Particles are approximately 90 x 60 nm. Diessner and Stedman, unpublished.

Following its discovery, SSV1 (then known simply as a virus-like particle or VLP) was instrumental in the work of Wolfram Zillig and colleagues that showed archaeal promoters are homologous to eukaryotic RNA polymerase II promoters (Reiter et al., 1988). As work with hyperthermophilic Archaea began to intensify, the need for genetic tools increased. Because SSV1 DNA is readily transformed into *Sulfolobus*, the virus efficiently spreads throughout a culture without apparent lysis, and cells cannot be cured of infection, there was interest in exploiting SSV1 as a genetic tool (Schleper et al., 1992). Stedman et al. (1999) were the first to show that SSV1 can stably maintain large insertions of

DNA (2.96 kb) without any obvious adverse effects. This work identified a number of sites in the SSV1 genome amenable to insertion, providing the first evidence for essential and non-essential SSV1 open reading frames (ORFs) as well as providing the first viral-based shuttle vectors for *Sulfolobus*. Shuttle vectors have since allowed for the creation of deletion mutants in SSV1 using long inverse PCR (LIPCR), expanding the toolbox of SSV genetics (Clore and Stedman 2006; Iverson and Stedman 2012; Chapters 2 and 3).

Comparative genomics of archaeal viruses

Bioinformatic techniques (e.g. BLAST) can provide valuable information about the function of uncharacterized proteins, helping to guide experimental studies (Prangishvili et al., 2006). Because these techniques depend on the presence of detectable sequence similarity, they are severely limited in their application when analyzing the virosphere where more than one-third of the ORFs lack recognizable homology to known ORFs in the public databases (Yin et al., 2008). This seems to be even more pronounced for viruses infecting the Archaea, where it is not uncommon for greater than 90% of the ORFs to lack recognizable homologs (Prangishvili et al., 2006). Structural genomics has proven to be a viable alternative, but is time consuming and inefficient (see below).

An exhaustive analysis of all crenarchaeal virus genomes was conducted using sensitive bioinformatic tools in an effort to tease out any homology that

could have been missed by previous studies (Prangishvili et al., 2006). While some new homologs were detected, the function of the majority of archaeal virus ORFs remained mysterious. Small ribbon-helix-helix (RHH) and helix-turn-helix (HTH) proteins were shown to be very common in crenarchaeal viruses, most of which are probably involved in transcription regulation of the virus and/or host. SSV1 was predicted to contain two RHH proteins, SSV1-C80 and SSV1-E51. A homolog of SSV1-E51 (SSV8-E73) was characterized and confirmed to contain the RHH structural motif (Schlenker et al., 2012). A putative HTH-containing protein was also predicted to be encoded in SSV1 (encoded by ORF *f93*) and the structure was later confirmed experimentally (Kraft et al., 2004a). Interestingly the product of ORF *f112* of SSV1 was shown to have a HTH motif yet was not predicted by Prangishvili et al., serving as a reminder that bioinformatic predictions are incomplete (Menon et al., 2008). Somewhat surprisingly, three putative SSV1 proteins (SSV1-B129, SSV1-A79, and SSV1-A45) are predicted to contain C_2H_2 Zn-finger (ZNF) structural motifs. C_2H_2 ZNF motifs are ubiquitous in eukaryotes but are not common in bacteria or Archaea (Prangishvili et al., 2006). However, this motif does seem to be common in the *Fuselloviridae* and other crenarchaeal viruses (Prangishvili et al., 2006).

A few other SSV1 ORF products have been assigned putative functions based on homology. The SSV1 integrase belongs to the tyrosine recombinase family of proteins and was initially the only SSV1 protein to which a function could be assigned on the basis of sequence alone (Palm et al., 1991;

Muskhelishvili et al., 1993; Serre et al., 2002). SSV1 ORF *d244* was predicted to encode a RecB-like endonuclease and a homolog (from SSV8 ORF *d212*) was later shown to exhibit a nuclease-like fold (Prangishvili et al., 2006; Menon et al., 2010). Finally, the product of SSV1 ORF *b251* was predicted to contain a NTP-binding motif characteristic of P-loop ATPases and exhibits limited similarity to the bacterial replication initiation protein DnaA (Koonin, 1992). A recent comprehensive bioinformatic analysis of crenarchaeal virus ATPases classifies SSV1-B251 and all fusellovirus homologs of this ORF as lon-like proteases (Happonen et al., 2014).

Comparative genomics of fuselloviruses

Since the publication of the SSV1 genome in 1992, eleven other fusellovirus-like genomes have been reported (Table 1-1). Comparative genomic analyses of fusellovirus genomes have been useful in determining the core fusellovirus genome (i.e. ORFs found in all fuselloviruses) as well as providing clues for the functions of some fusellovirus proteins (Redder et al., 2009; Held and Whitaker 2009; Figure 1-3 and Table 1-2). Additionally, the more fusellovirus genomes that are sequenced will enable the identification of ORFs that are likely bonafide protein coding genes (Prangishvili et al., 2006).

All fusellovirus genomes are double-stranded, circular DNA approximately 14.5 – 17.5 kb in length, with the exception of the ~24 kb genome of ASV1 (Redder et al., 2009; Table 1-1). Fusellovirus genomes exhibit an interesting

dichotomy: one half of the genome contains highly conserved ORFs while the other half exhibits significantly higher variation (Figure 1-3). Nucleotide sequence similarity between fuselloviruses generally occurs in small regions of the genome, typically within conserved ORFs (Redder et al., 2009). The fusellovirus "core" genome has steadily decreased as more fusellovirus genomes have been sequenced (Stedman et al., 2003; Wiedenheft et al., 2004; Redder et al., 2009). A number of putative transcription factors are encoded in the region of the fusellovirus genome corresponding to the conserved T5 and T6 "early" transcripts, implicating them in regulation of viral and/or host transcription as well as viral DNA replication (Fröls et al., 2007; Redder et al., 2009). All fuselloviruses encode a tyrosine recombinase that is responsible for viral integration into a host tRNA gene (Reiter et al., 1989). As expected, all fuselloviruses harbor well-conserved VP1 and VP3 homologues, however, only SSV1, SSV6, ASV1, and SMF1 encode a VP2 homolog (Redder et al., 2009; Servin-Garcidueñas et al., 2013).

Discovery of SSV6 and ASV1 led to the identification of putative tail-fiber encoding modules in the *Fuselloviridae* (Redder et al., 2009). The ASV1 and SSV6 putative tail fiber module is clearly distinct from the other fuselloviruses and corresponds to a morphologically distinct tail structure and virions that are not prone to forming rosettes. Homologues of SSV1-A153 and SSV1-B251 are present in every fusellovirus and surprisingly in the satellite virus pSSVx, leading to the hypothesis that these ORFs are involved in packaging (Arnold et al.,

1999). The discovery of the satellite pSSVi has cast doubt on this hypothesis as pSSVi is likewise packaged into fusellovirus-like particles but does not encode SSV1-B251 or SSV1-A153 homologs (Wang et al., 2007; Figure 1-3).

Table 1-1: List of fusellovirus genomes

Virus	Genome Size (bp)	# of ORFs	Accession	Reference
SSV1	15,465	35	NC_001338	Palm et al., 1991
SSV2	14,795	35	NC_005265	Stedman et al., 2003
SSV3	15,230	32	n/a	Stedman et al., 2006
SSV4	15,135	34	NC_009986	Peng 2008
SSV5	15,330	34	NC_011217	Redder et al., 2009
SSV6	15,684	33	NC_013587	Redder et al., 2009
SSV7	17,602	33	NC_013588	Redder et al., 2009
SSV8 [A]	16,473	37	NC_005360	Wiedenheft et al., 2004
SSV9 [B]	17,385	31	NC_005361	Wiedenheft et al., 2004
SSVL	14,461	31	n/a	Personal communication
ASV1	24,186	38	NC_013585	Redder et al., 2009
SMF1 [C]	14,847	24	NC_020882	Servin-Garcidueñas et al., 2013

[A] SSV8 formerly referred to as *Sulfolobus* spindle-shaped virus Ragged Hills
[B] SSV9 formerly referred to as *Sulfolobus* spindle-shaped virus Kamchatka
[C] The genome sequence of SMF1 was constructed from a metagenome, a viral particle has not been isolated

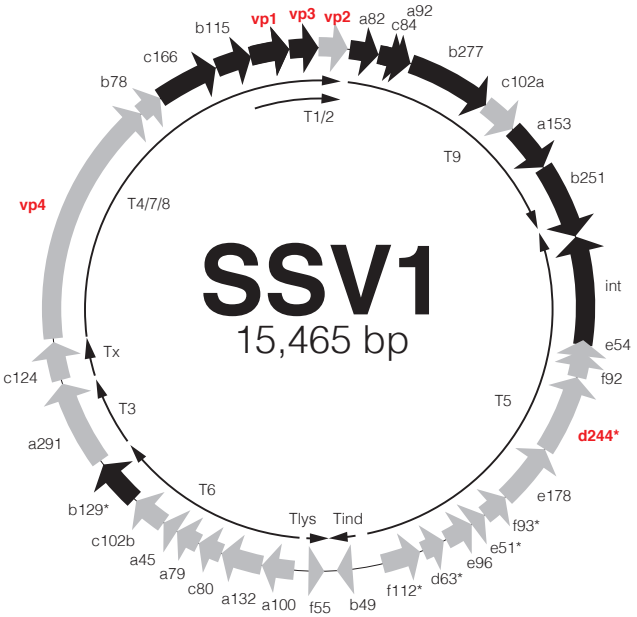


Figure 1-3: Map of the SSV1 genome.
Block arrows denote SSV1 ORFs and are labeled as in Palm et al., 1992 and Fusco et al., 2013. Black filled ORFs denote those that are universally conserved and belong to the fusellovirus core (Table 1-2), grey-filled ORFs are not universally conserved. Thin arrows on interior indicate known SSV1 viral transcripts (Reiter et al., 1987; Fröls et al., 2007; Fusco et al., 2013). ORF labels in red indicate proteins identified in purified SSV1 virions (Menon et al., 2008; Quemin et al., 2015). Labels with an asterisk (*) denote ORFs that have been structurally characterized (See text for references). The eleven fusellovirus genomes used for this analysis: SSV1, SSV2, SSV3, SSV4, SSV5, SSV6, SSV7, SSV8, SSV9, SSVL, and ASV1 (Table 1-1).

Table 1-2: Open reading frames comprising the fusellovirus core

SSV1 ORF	Annotation	Reference
VP1	Major capsid protein	Reiter et al., 1987
VP3	Minor capsid protein	Reiter et al., 1987
Integrase	Integrase	Mushkelishvili et al., 1993
A153	Unknown	N/A
A82	Unknown	N/A
A92	Unknown	N/A
B115	Putative HTH transcription regulator	Prangishvili et al., 2006
B129	Putative C_2H_2 ZNF transcription regulator	Lawrence et al., 2009
B251	ATPase, DnaA homologue, lon-like protease	Koonin 1992
B277	Unknown	N/A
C84	Unknown	N/A
C166	Unknown	N/A

❖ SSV1 ORFs named as in Palm et al., 1992. Fusellovirus core ORFs determined by comparing the proteins from 11 fusellovirus genomes (Table 1-1). SMF1 was omitted from this survey as it is unclear if it is a complete fusellovirus genome (Servin-Garcidueñas et al., 2013).

Archaeal virus protein structure and function

Because the vast majority of archaeal virus proteins exhibit little-to-no homology to known sequences, structural studies have been initiated on the basis that the three-dimensional structure of a protein is conserved on an evolutionary time scale and has the potential to reveal homologs that would otherwise have gone undetected (Lawrence et al., 2009; Dellas et al., 2013). The results of over ten years of structural studies on archaeal viruses has shown that the vast majority of solved structures (~80%) conform to previously known

protein folds, despite practically non-existent sequence similarity. This trend holds true for the fuselloviruses, from which nine structures have been solved (Lawrence et al., 2009; Dellas et al., 2013).

Archaeal viruses, including fuselloviruses, seem to encode a large number of DNA binding proteins that may be involved in transcription regulation (Prangishvili et al., 2006). The winged helix-turn-helix (wHTH) motif is present in the crystal structures of the protein products of ORFs *f93* and *f112* from SSV1 and the crystal structure of STIV1-F93 (Kraft et al. 2004a; Menon et al., 2008; Larson et al., 2007a). SSV1-F93 was shown to exist as a dimer in solution and is hypothesized to recognize a palindromic or pseudo-palindromic DNA sequence on the host chromosome, as no binding site in the SSV1 genome could be identified (Kraft et al., 2004). Interestingly, a well-conserved motif (LTEKG) is found near the N-terminus of all three fusellovirus F93 homologs (SSV1-F93, SSV9-E81 and ASV1-95) and also in STIV1-F93. This patch appears to be distant from the DNA-binding interface and could be involved in interactions with another protein, most likely host-derived (Larson et al., 2007a; Appendix B). The product of SSV1 ORF *f112* unexpectedly crystallizes as a monomer and apparently also exists as a monomer in solution, although it is possible the disordered N-terminus of SSV1-F112 may be involved in interactions with an unidentified protein to form a heterodimer (Menon et al., 2008).

The solution structure of the product of ORF *e73* from SSV8 has been solved and displays a RHH motif (Schlenker et al., 2012). RHH motifs have not

been observed in eukaryotes but seem to be very common in Bacteria, Archaea, and their viruses (Schreiter et al., 2007). SSV8-E73 is well conserved within the *Fuselloviridae*, with homologs found in six of the 12 published fusellovirus genomes (Appendix B). However, the SSV8-E73 homolog in SSV1 (SSV1-E51) lacks the C-terminal extension found in other fusellovirus homologues. SSV8-E73, like almost all RHH proteins, exists as a dimer in solution. The DNA-binding site of SSV8-E73 has not been identified. However, SSV8-E73 does exhibit non-specific DNA-binding that can be completely abolished by a K11E mutation in the antiparallel Beta-sheet predicted to make base-specific contacts with the DNA (Schlenker et al., 2012). It should be noted that this mutation does not affect the structure of the protein. In addition to the product of SSV1 ORF *e51*, the products of SSV1 ORFs *c80* and *f55* are also predicted to display RHH structural motifs (Prangishvili et al., 2006; Fusco et al., 2013).

Several fusellovirus proteins are predicted to or have been shown to encode one or more zinc finger motifs (ZNF) (Prangishvili et al., 2006). SSV1 encodes three putative proteins containing the ZNF motif (SSV1-B129, SSV1-A79, and SSV1-A45), all of which occupy the same T6 “early” transcript (Reiter et al., 1987b; Prangishvili et al., 2006; Fröls et al., 2007). The crystal structure of the product of the universally conserved SSV1 ORF *b129* has been solved and shown to contain tandem C_2H_2-type ZNF motifs (Lawrence et al., 2009). Additionally, SSV1-B129 has been shown experimentally to nonspecifically bind dsDNA (Lawrence, personal communication). The protein AFV1p06 from

Acidianus filamentous virus 1 (AFV1) also contains a C_2H_2-like ZNF motif and has homologs in seven of the 12 published fusellovirus genomes, including SSV1-A45 (Guillere et al., 2013). AFV1p06 and fusellovirus homologues encode a well-conserved hydrophobic C-terminal extension that is distant from the putative DNA-binding interface that is hypothesized to be involved in protein-protein interactions that may play a regulatory role (Guillere et al., 2013).

Structures of several non-DNA-binding proteins have also been reported. The crystal structure of the catalytic domain of the SSV1 integrase, the first archaeal integrase structure solved, yielded insights into the mechanism of fusellovirus integration (Eilers et al., 2012). The crystal structure of the product of ORF *d212* from SSV8, homologous to SSV1-D244, revealed a new member of the PD-(D/E)XK nuclease superfamily and shares significant structural homology to a holiday junction resolvase from *S. solfataricus* (Menon et al., 2010). The crystal structure of SSV1-D63 has also been solved and contains a dimeric 4-helix bundle motif that is characteristic of a multitude of proteins with diverse functions (Kraft et al., 2004b). The wide-spread and divergent applications of this fold has made a functional prediction impossible. Comparison of SSV1-D63 homologs from all fuselloviruses reveals several highly conserved surface residues and implies that SSV1-D63 might be interacting with another macromolecule (Kraft et al., 2004b; Appendix B).

The structure-based approach exemplified above does not always yield recognizable protein folds. For example, the crystal structure of the product of

ORF SSV1 *e96* yields no significant homologs via structure comparisons, suggesting that SSV1-E96 exhibits a novel fold (Lawrence et al., 2009). Curiously, SSV1 is the only fusellovirus to encode an ORF *e96* ortholog. Other examples of novel protein folds from crenarchaeal viruses include AFV1-109, AFV1-99 and its homolog in STIV1, STIV1-B116 (Keller et al., 2007; Goulet et al., 2009; Larson et al., 2007b).

Gene expression in viruses of Sulfolobus

The last decade has seen a number of studies focused on elucidating the transcriptional response to virus infection and induction for several hyperthermophilic archaeal viruses and hosts. These studies have helped to elucidate archaeal virus transcription patterns, or lack thereof, as well the effects of virus infection on the host RNA expression. Furthermore, transcript analysis has helped to hypothesize or confirm roles for some of the multitude of uncharacterized genes in archaeal virus genomes.

It has been known for some time that SSV1 replication could be induced by UV light (Martin et al., 1984) and this characteristic was exploited by Reiter et al. (1987b) and others to study transcription in SSV1. Reiter's initial findings were later confirmed and expanded upon using cDNA microarrays (Fröls et al., 2007; Figure 1-3). Exposure of an SSV1 lysogen (defined herein as a host cell harboring SSV1 in the carrier state) to UV light revealed that SSV1 transcription was temporally regulated and begins with expression of the "immediate-early"

T_{ind} transcript, followed by three "early" transcripts (T5, T6, and T9) and four "late" transcripts (T1/2, T3, and T4/7/8) (Reiter et al., 1987b; Figure 1-3). The entire transcription cycle after UV-induction is completed within 8.5 hr. The timing of the SSV1 transcripts correlates well with predicted SSV1 gene functions, i.e. structural genes are encoded by the late transcripts and putative transcription factors are encoded by the early transcripts. The T9 transcript is the last up-regulated transcript prior to DNA replication, implicating the gene products in viral replication. The late monocistronic transcript Tx was previously unidentified and the timing of its expression hints at a possible structural role for ORF *C124*, although it has not been identified in purified SSV1 virions. Because SSV1 is apparently the only strongly UV-inducible fusellovirus, it is unclear how relevant these data are to the *Fuselloviridae* in general.

In un-induced SSV1-infected cells, incomplete versions of the "early" transcripts but complete "late" transcripts were continuously present (Reiter et al., 1987b; Fröls et al., 2007). Although the upstream region of the T5 transcript was not observed, the region of T5 containing the integrase gene was continuously observed and suggests the integrase gene may have its own transcript nestled within the T5 transcript (Clore thesis 2008). It should also be noted that the immediate early T_{ind} transcript was not observed in un-induced cells, implying it (and the encoded *b49* ORF) is only active early in the viral response to UV irradiation. A recent study aimed at identifying genes involved in the maintenance of SSV1 lysogeny identified a novel transcript that is continuously

expressed throughout lysogeny and encodes a 55 residue protein (F55) that was shown to bind weakly (Kd ~1-9 μM) and with differential affinity to the T5, T6, T_{ind}, and T_{lys} promoters (Fusco et al., 2013). Because F55 binding overlaps with BRE or TSS promoter elements, it is hypothesized that F55 plays a negative regulatory role, repressing transcription from these genes until a signal (e.g. UV light) induces virus replication. A recent follow up study reported F55 binding to these promoters *in vivo* and that binding dissipates shortly after UV induction (Fusco et al., 2015b). This hypothesis fits well with the absence of T5, T6 and T_{ind} transcripts throughout lysogeny.

The only other fusellovirus to have its transcripts analyzed is the closely related SSV2 (Ren et al., 2013). *S. solfataricus* P2 cells with and without the satellite virus pSSVi were infected with SSV2 at an unknown multiplicity of infection (MOI) and the viral and host transcription was measured using cDNA microarrays (Ren et al., 2013). Similar to SSV1, SSV2 transcription is temporally regulated and genes can be classified as “early” or “late” depending on whether they are expressed before or after 4.5 hours post infection (h.p.i.). The ORF *305* is the first SSV2 ORF expressed after infection and is followed shortly thereafter by the capsid protein genes *vp1* and *vp3*. SSV2 ORF *305* is a reported homolog of SSV1 ORF *a291* which is constitutively expressed in SSV1-infected cells and up-regulated late following UV induction (Fusco et al., 2013; Fröls et al., 2007). The presence of pSSVi was shown to delay viral DNA replication as well as alter the timing, but not the chronology, of SSV2 transcript production. It should be

noted that because these experiments were performed at an unknown MOI it cannot be confirmed if a synchronous infection was achieved.

Transcript studies of fuselloviruses have focused on gene expression immediately following infection or induction but not much work had been done on gene expression during chronic infection until recently (Fusco et al., 2013; Fusco et al., 2015c). In these works, *S. solfataricus* P2 derivative LnF1 was infected with SSV1 or SSV2 at an unstated MOI and expression of viral genes was measured. Interestingly, both viruses express several homologous genes (*vp1*, *vp3*, SSV1-*a291*/SSV2-*305*, and SSV1-*c124*/SSV2-*126*) that have been hypothesized as the minimum set of genes required for the maintenance of chronic infection (Fusco et al., 2015c). Despite the similar expression of transcription, chronic infection with SSV1 and SSV2 had dramatic differences on host gene expression (Fusco et al., 2015c). While SSV1 hardly perturbed expression in the host, SSV2 infection elicited a more dramatic reaction from the host that was reminiscent of the effects caused by infection with lytic phages SIRV2 and STIV. Intriguingly, when LnF1 cells were co-infected with both SSV1 and SSV2 the host transcriptome resembled that of cells infected with SSV1, suggesting SSV1 somehow mitigates the effect of the SSV2 (Fusco et al., 2015c).

Two other unrelated hyperthermophilic archaeal viruses infecting *Sulfolobus* have also had their transcripts analyzed. Expression of RNAs by the lytic rudiviruses SIRV1 and SIRV2 appeared to exhibit little to no temporal regulation,

as transcription initiation occurs from 20-30 different sites and results in expression of almost all viral genes within 30 minutes of infection (Kessler et al., 2004). Interestingly, a later study of SIRV2 infection on a highly susceptible *S. solfataricus* P2 derivative (harbors a large CRISPR deletion) revealed that SIRV2 transcription is more temporally regulated than previously thought (Okutan et al., 2013). Analysis of transcription in the lytic virus STIV likewise reveals minor temporal regulation and a slower transcription cycle (~32 hr.) compared to the fuselloviruses and rudiviruses (Ortmann et al., 2008). STIV transcripts were first identified 8 hours post infection (h.p.i). and most genes are significantly expressed within 16 h.p.i. The STIV structural genes are expressed late in the transcription cycle (16-24 h.p.i) albeit in different amounts relative to each other. The available data imply that the absence of temporal regulation may be characteristic of a lytic life cycle, whereas non-lytic archaeal viruses may need to exhibit tighter control on the timing of transcription (Ortmann et al., 2008; Okutan et al., 2013; Kessler et al., 2004). It has been suggested that infection by lytic archaeal viruses results in much more pronounced changes in the host transcriptome compared to non-lytic infections, although recent work with the non-lytic virus SSV2 questions this hypothesis (Quax et al., 2013; Fusco et al., 2015c). Clearly more experimental data from both lytic and non-lytic archaeal viruses is needed.

DNA replication in hyperthermophilic Archaea and viruses:

Very little is known about how hyperthermophilic archaeal viruses replicate their DNA. Almost all isolated and sequenced hyperthermophilic archaeal viruses do not encode a recognizable DNA polymerase (Peng et al., 2007; Wang et al., 2015). Analysis of host transcripts following infection with STIV and SIRV2 show up-regulation in many host genes involved in DNA replication and repair, suggesting viral dependence on these enzymes (Ortmann et al., 2008; Kessler et al., 2004). Because viral origins of replication appear to be distinct from those of the host (and other viruses), it is unclear how the replication machinery is recruited (Wang et al., 2015). Furthermore, the presence of linear and circular dsDNA viral genomes suggests there are likely to be multiple varied DNA replication strategies.

The eukaryotic DNA replication machinery is strikingly similar to the machinery used by the Archaea, although the archaeal machinery is less complex. This reduced complexity makes the Archaea, and the hyperthermophilic Archaea in particular, a good model organism for the study of eukaryotic DNA replication. DNA Replication in *S. solfataricus* has been well studied and many of the components are well understood (Duggin and Bell, 2007). *S. solfataricus* contains three origins of replication, at which a number proteins are assembled to yield the replicative complex. The components of the *S. solfataricus* replisome includes single stranded binding proteins (RPA), replicative helicase (MCM), primase, DNA polymerase (PolB1), heterotrimeric

sliding clamp (PCNA-1, PCNA-2, PCNA-3), and the clamp loader (Bell et al., 2003; Kelman and Kelman 2014)

The best understood examples of archaeal virus DNA replication come from the rudiviruses SIRV1 and SIRV2. Initial studies indicated head-head and tail-tail replication intermediates, which combined with similarities to *Poxviridae* genome arrangement, led to a proposed replication model where a Rep protein nicks the viral DNA at a specific site yielding a 5'-DNA adduct and a free 3'-hydroxyl that can prime DNA synthesis (Peng et al., 2001). A dimeric rep protein from SIRV1 was later purified and found to display nicking activity at the predicted site within viral genome and is the first example of a Rep protein implicated in replication that does not undergo rolling circle or rolling hairpin replication (Oke et al., 2011). A yeast two-hybrid approach identified five SIRV2 proteins that interact with the host PCNA clamp, providing some clues into how recruitment of the host replication machinery might occur (Gardner et al., 2014).

The situation is even more nebulous in the *Fuselloviridae*. SSV1 transcription data suggests the ORFs encoded on the T9 transcript are likely involved in DNA replication as they are expressed just prior to an observed increase in SSV1 copy number and almost all of these belong to the fusellovirus core (Fröls et al., 2007; Figure 1-3 and Table 1-2). SSV1 ORF *b251* is predicted to be homologous to the bacterial replication initiation protein DnaA and is conserved among the *Fuselloviridae* (Koonin et al., 1992; Appendix A). The putative origin of replication in SSV1 has been hypothesized to lie near the T_{ind}

transcript but this remains to be confirmed (Fröls et al., 2007; Cannio et al., 1998). SSV9, SSV7, and ASV1 encode homologs of yeast Rad3 DNA helicase, the role of which is unclear in the viral life cycle (Appendix A).

CRISPR-Cas systems: overview and interactions with archaeal viruses:

Originally discovered in 1987, research on prokaryotic CRISPR (clustered regularly interspaced short palindromic repeats) -mediated immune systems has exploded over the last 5-10 years. CRISPR loci and Cas (CRISPR-associated) genes are present in ~50% of bacterial and ~85% of archaeal genomes and have been observed in some cases to function as a prokaryotic adaptive immune system against viruses and other mobile genetic elements (Grissa et al., 2007). There are currently three recognized types of CRISPR-Cas systems (I, II, and III), and each is further divided into multiple subtypes (Sorek et al., 2013; Barrangou and Marraffini 2014). Only two *cas* genes are universal to each CRISPR-Cas system (Cas1 and Cas2) while the remaining *cas* genes vary from system to system (Barrangou and Marraffini 2014). A CRISPR locus is comprised of many short repeat elements separated by unique sequences called "spacers", at least some of which are derived from invading plasmids and viruses (Sorek et al., 2013). Environmental data show a high degree of variability in the spacer content of natural populations, where it's rare for the most recently acquired spacers to be identical between two cells (Andersson and Banfield 2008). Furthermore, only the most recently acquired spacers match virus

sequences suggesting that the viruses are under constant pressure to mutate in order to avoid host defenses (Andersson and Banfield 2008; Sun et al., 2013; Banfield et al., 2015).

The CRISPR-mediated immune response is divided into three stages: adaptation, expression, and interference (Sorek et al., 2013). Each stage utilizes a different set of associated enzymes and these enzymes vary among the many CRISPR types/subtypes. Adaptation is the process by which new viral and plasmid sequences are incorporated into the CRISPR repeat-spacer array and is the least understood of the three stages (Erdmann and Garrett 2012). Newly incorporated spacers seem to be added to the leader-proximal region of the CRISPR array, however there is experimental evidence suggesting this may not be universal to all CRISPR systems (Erdmann et al., 2013). During the expression stage, the entire CRISPR locus is transcribed as a large precursor RNA (pre-crRNA) that is processed by Cas proteins into individual crRNAs, each containing a repeat and spacer sequence. CrRNAs are individually recruited by a large multi-protein "interference" complex. The crRNA guides the interference complex to the foreign nucleic acid and facilitates degradation of the invader.

Sulfolobus possesses complex CRISPR-Cas systems and includes some of the best experimental models in the field (Zhang et al., 2013; Manica et al., 2013). CRISPR-mediated resistance has been demonstrated in *Sulfolobus* for a few viruses and plasmids while a number of Cas proteins and complexes have been thoroughly studied (Manica et al., 2013; Erdmann et al., 2013; Erdmann et

al., 2014a; Manica et al., 2011). Transcript analyses following infection with crenarchaeal viruses STIV and SIRV2 revealed a robust host CRISPR-Cas response (Quax et al., 2013; Maaty et al., 2012). It's unclear if the dramatic cellular response results from infection with lytic viruses, which are a major threat to cell mortality, or if this is a more general response of *Sulfolobus* to any virus.

Work on *S. solfataricus* strain P2 and *S. islandicus* REY15A showed that mixtures of viruses, which cells presumably encounter in nature, elicit different CRISPR immune responses compared to infections with single viral isolates (Erdmann et al., 2012; Erdmann et al., 2014). Spacer uptake in P2 could only occur in the presence of *Sulfolobus* monocaudovirus 1 (SMV1) and a co-infecting virus or plasmid (STSV2 or pMGB1, respectively) and these spacers were shown to confer immunity against the matching genetic element (Erdmann et al., 2012). Interestingly, spacers were only acquired for the co-infecting genetic element and never for SMV1 despite the continued presence of SMV1 in the culture, indicating that SMV1 somehow avoids the host CRISPR adaptation machinery while simultaneously activating the spacer uptake from other genetic elements. A later study in *S. islandicus* REY15A confirmed these results and expanded upon them (Erdmann et al., 2014). They were able to show SMV1 spacer acquisition could be accomplished only following "cold-shock" of the cells and that spacer acquisition seems to occur simultaneously with SMV1 replication. However, these spacers do not confer immunity to SMV1 suggesting that SMV1 evades the host interference machinery through an unknown mechanism.

While a significant amount of CRISPR research has been done in *Sulfolobus*, the majority of this work has not been focused on fuselloviruses. A screen of the CRISPR arrays from several *Sulfolobus*, *Acidianus, and Metallosphaera* genomes revealed a large number of spacer sequences matching known crenarchaeal viruses and a significant proportion of these match fuselloviruses (Shah et al., 2009). Recombinant SSV1 and *S. solfataricus* P2 were used in one of the first demonstrations of CRISPR-mediated defense against an invading virus (Manica et al., 2011). Because *S. solfataricus* P2 does not naturally harbor any spacer matches to SSV1, a protospacer to pNOB8 was inserted into SSV1 and this recombinant virus was shown to be significantly less infectious than wild-type SSV1. Infection of *S. solfataricus* P2 derivative LnF1 with SSV1 failed to activate the CRISPR system, but infection with SSV2 did result in CRISPR activation and was very reminiscent of the host response to infection with lytic viruses STIV and SIRV2 (Fusco et al., 2015c). However, when LnF1 cells were doubly-infected with SSV1 and SSV2, CRISPR activation was reportedly absent and host gene expression was similar to that of the SSV1 singly-infected cells. These data are intriguing and suggest that SSV1 somehow mitigates the host response to infection.

The observation that most loci for cellular immunity, including CRISPR-Cas, also encode putative toxin genes led to an intriguing hypothesis for the incorporation of the latter into the host immune response (Makarova et al., 2012). These theoretical toxin genes are expressed simultaneously with other *cas*

genes during the CRISPR response and may induce dormancy, allowing the cell to 'buy time' to mount a sufficient immune response. If the immune response fails to contain the infection, dormancy continues and results in cell death thus preventing the spread of infection to the rest of the population. Another intriguing aspect of this hypothesis is that it provides a temporal mechanism by which a cell can acquire de novo spacers against a novel genetic element and mount a CRISPR-mediated immune response, a process that remains obscure. A well conserved *cas* gene (Csa5) was shown to be highly toxic when overexpressed in *S. solfataricus*, supporting the proposed toxicity of some *cas* genes (He et al., 2014). It should be noted that the most highly expressed host gene following infection with STIV is a gene of unknown function but homology searches suggest it may be a toxin (Quax et al., 2013; Ortmann et al., 2008). Infection with *S. islandicus* cells with SSV9 at low MOI was shown to induce dormancy and in some cases cell death (Bautista et al., 2015). Infection of cells harboring a 100% spacer match to SSV9 resulted in a 24-48 hr. growth delay from which the cells recovered and successfully eliminated the virus. However, cells harboring inactivated CRISPR systems were not able to clear the virus and never recovered from the initial growth retardation. The results from this study fit well with the induced dormancy hypothesis proposed by Makarova et al. but it is clear that more data is needed to elucidate the observed phenomena.

Conclusion:

The significant amount research performed in the past few decades illustrates the astonishing diversity possessed by the viruses infecting the Archaea. In particular, viruses infecting the Crenarchaea exhibit a multitude of bizarre morphologies that seem be exclusive to the archaeal virosphere and probably have much teach us about protein folding and stability in extreme environments. Furthermore, the genetic diversity possessed by these viruses is at the same time both fascinating and confounding. The *Fuselloviridae* are one of the best studied families of archaeal viruses and have proven to be very useful subjects for the study of archaeal virology.

As detailed above, very little is known about the functions of the vast majority of fusellovirus proteins. To address this, we have completed a comprehensive genetic analysis of SSV1 by constructing deletion and/or insertion mutants in every ORF of the SSV1 genome. Mutagenesis studies are useful as they can provide clues for the function of mysterious proteins. The results of this work dramatically expand on previous knowledge of the genetic requirements of SSV1 and provide further insight into the function of several SSV1 proteins.

We have also initiated biochemical analyses of SSV1 proteins in an effort to better characterize their function. The DNA binding protein VP2 was identified in purified SSV1 virions and is believed to reside within the capsid, bound to the viral DNA (Reiter et al., 1987). However, the specifics of VP2 DNA binding and

its physiological significance is not well understood. To address these questions, we have developed a robust protocol for the purification of recombinant VP2 protein from *E. coli*. Our results suggest that VP2 exists as a monomer in solution and appears to bind non-specifically to dsDNA. This work should be useful for follow-up studies of the three-dimensional structure of VP2 and a more in depth characterization of its DNA binding affinity.

Finally, we present the first insights into the kinetics of a fusellovirus infection by performing one-step growth curve experiments with SSV1. These experiments have been challenging to perform with SSV1 due to difficulties in isolating the large number of virus required. We show that low MOI infections with SSV1 and SSV1 mutants result in the production of high viral titers which can subsequently be used to perform one-step growth curve experiments at high MOIs. The establishment of a reliable protocol for a one-step growth curve will be invaluable for further characterization of the numerous SSV1 mutants we have isolated.

Chapter 2: A genetic study of SSV1, the prototypical fusellovirus

This chapter is based on the following publication:

Eric Iverson and Kenneth M. Stedman. "A genetic study of SSV1, the prototypical fusellovirus"

Front Microbio. 3:200. doi: 10.3389/fmicb.2012.00200.

Abstract

Viruses of thermophilic Archaea are unique in both their structures and genomic sequences. The most widespread and arguably best studied are the lemon-shaped fuselloviruses. The spindle-shaped virus morphology is unique to Archaea but widespread therein. The best studied fusellovirus is SSV1 from Beppu, Japan, which infects *Sulfolobus solfataricus*. Very little is known about the function of the genes in the SSV1 genome. Recently we have developed genetic tools to analyze these genes. In this study, we have deleted three SSV1 open reading frames (ORFs) ranging from completely conserved to poorly conserved: *vp2*, *d244*, and *b129*. Deletion of the universally conserved ORF *b129*, which encodes a predicted transcriptional regulator, results in loss of infectivity. Deletion of the poorly conserved predicted DNA-binding protein gene *vp2* yields viable virus that is indistinguishable from wild-type. Deletion of the well-conserved ORF *d244* that encodes a predicted nuclease yields viable virus. However, infection of *S. solfataricus* with virus lacking ORF *d244* dramatically retards host growth, compared to the wild-type virus.

Introduction

Viruses of Archaea are very poorly understood with only about 50 known archaeal viruses relative to the ca. 5000 characterized viruses of bacteria, plants, and animals (Pina et al., 2011). The best studied of archaeal viruses are those infecting the thermoacidophiles, with an unprecedented new seven virus families introduced in the last few years to accommodate the astonishing morphological and sequence diversity present in these viruses (Pina et al., 2011).

The *Sulfolobus* spindle-shaped viruses (SSVs) of the family *Fuselloviridae* were the first discovered and probably the best studied family of archaeal viruses. SSVs are found throughout the world in high temperature (>70˚C) and acidic (pH < 4) environments where their hosts, *Sulfolobus solfataricus* and its close relatives thrive (Wiedenheft et al., 2004; Held and Whitaker, 2009). The type virus, SSV1, encodes a positively supercoiled, 15.5 kbp circular dsDNA genome (NC_001338.1) that is enclosed within a lemon or spindle-shaped capsid (Yeats et al., 1982; Martin et al., 1984; Nadal et al., 1986). The genome encodes 34 open reading frames (ORFs; Palm et al., 1991), most of which have no recognizable homologs apart from other *Fuselloviridae*. The only SSV1 gene with clear homology to proteins outside the *Fuselloviridae* is the viral *integrase*, encoded by ORF *d355*. The main structural proteins purified from virus particles are the major and minor capsid proteins *vp1* and *vp3* and the putative DNA packaging protein *vp2* (Reiter et al., 1987a). More recently, mass spectrometric analysis of SSV1 virions revealed two additional proteins: the products of ORFs

c792 and *d244* (Menon et al., 2008; Figure 2-1).

In the absence of homologous sequences, three complementary approaches have been used to try and determine the function of the proteins encoded in the SSV1 genome; structural genomics, comparative genomics, and genetics. Atomic resolution structures have been obtained by C. Martin Lawrence and his group for proteins encoded by SSV1 ORFs *b129*, *f112*, *d63*, *e96*, *f93*, and *d244* or their homologs from other fuselloviruses. The products of ORFs *b129*, *f112*, and *f93* resemble transcriptional regulators and *d244* a novel nuclease (Lawrence et al., 2009; Menon et al., 2010). However, the function of these proteins in virus replication remains to be determined. Two of these ORFs, *b129* and *d244*, are the targets of the current study.

In parallel, we and others have undertaken comparative genomic studies. Fifteen ORFs are completely conserved in 12 canonical SSV genomes (Stedman et al., 2003; Wiedenheft et al., 2004; Held and Whitaker, 2009; Redder et al., 2009; Stedman, unpublished; Figure 2-1). Most of the universally conserved genes are clustered in half of the genome with the notable exception of the *vp2* gene, a target of this study. Conservation in the rest of the genome is lower. Nonetheless, there are very few completely unique genes in the SSV1 genome (Figure 2-1). It is highly probable that the conserved genes are required for virus function, but again this has not been confirmed.

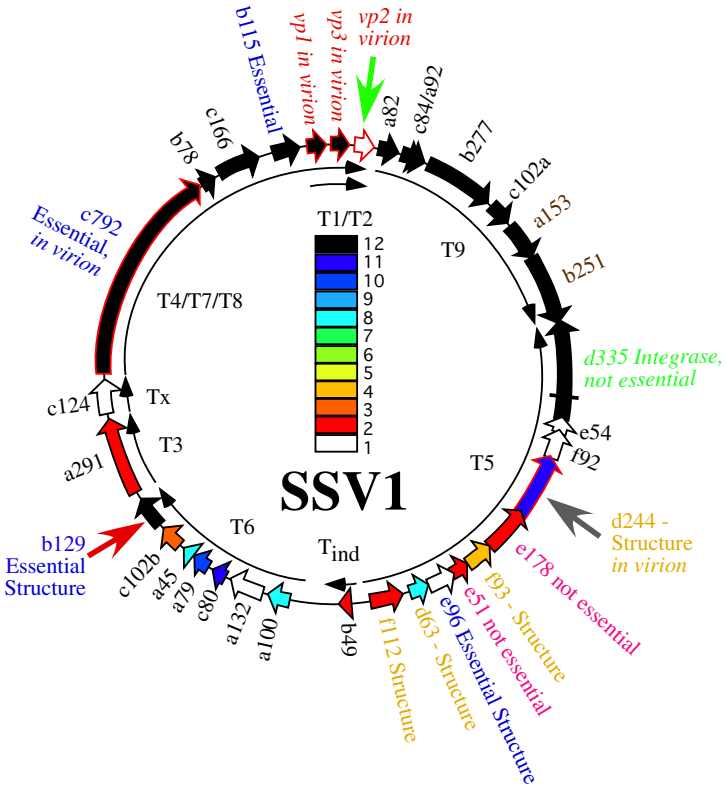


Figure 2-1: Genome map of SSV1
Open reading frames are shown as block arrows and labeled as in Palm et al. (1991). Virus structural protein genes (Reiter et al., 1987a) and other proteins found in the virion (Menon et al., 2008) are outlined in red and labeled as "in virion." Conservation of open reading frames in 12 canonical SSV genomes (SSV1, SSV2, SSV3, SSV4, SSV5, SSVRH, SSVK1, SSVL, SSVKM1, SSVKU1, SSVL2, and SSVGV1; Redder et al., 2009; Held and Whitaker, 2009; Stedman, unpublished) is listed with the color code in the middle of the genome with ORFs conserved in 12 genomes in black, ORFs conserved in 11 genomes in dark blue, etc. ORFs which did not tolerate insertion of the pBluescript plasmid are labeled as "Essential" in blue type. ORFs allowing insertion of the pBluescript plasmid without loss of virus function are labeled as "not essential" (Stedman et al., 1999). All ORFs whose products have been crystallized and structure determined are labeled as "Structure" (Lawrence et al., 2009; Menon et al., 2010). The gene for the SSV1-integrase is labeled in green and was shown to be not essential by deletion (Clore and Stedman, 2006). Transcripts are labeled as curved thin arrows (Reiter et al., 1987b; Fröls et al., 2007). ORFs targeted in this study are indicated with large arrows outside the genome map.

We developed methods for gene disruption in order to determine the requirements for genes in the virus genome directly. About 10 years ago, we showed that four SSV1 ORFs did not tolerate insertion of the 3.2 kbp pBluescript plasmid and allow virus function. Twelve other SSV1 ORFs appeared, indirectly,

to not tolerate insertion. However, two ORFs, *e178* and *e51*, were able to tolerate insertion of the entire pBluescript plasmid (Stedman et al., 1999). This result allowed the development of viral shuttle vectors and the beginnings of *Sulfolobus* genetics (Jonuscheit et al., 2003). Insertion of the pBluescript plasmid and up to ca. 5 kbp of exogenous DNA in these ORFs does not appear to have a noticeable effect on virus function (Stedman et al., 1999; Jonuscheit et al., 2003; Clore and Stedman, 2006; Albers et al., 2006).

However, insertion of large DNA fragments into the SSV1 genome is not straightforward and the possible insertion locations are limited. Therefore, Long Inverse PCR (LIPCR) using high-fidelity highly processive DNA polymerases (e.g., Phusion^{TM}) was developed to specifically change the SSV1 genome at single nucleotide resolution. LIPCR was used to delete precisely the SSV1 viral *integrase* gene. Surprisingly, this "integrase-less" SSV1 was functional (Clore and Stedman, 2006). However, consistent with its conservation, the virus lacking the integrase gene is at a competitive disadvantage relative to integrase-containing viruses (Clore and Stedman, 2006). All of the SSV1 ORFs that can be deleted or tolerate insertion without abrogating virus function are in the "early" transcript, T5, that is induced soon after UV-irradiation of SSV-infected cultures (Reiter et al., 1987b; Fröls et al., 2007).

Three ORFs in the SSV1 genome were targeted for gene disruption in this study. The *vp2* gene (NP_039802.1) was chosen for disruption because it is only present in SSV1 and the very distantly related SSV6 (Held and Whitaker, 2009;

Redder et al., 2009), and is in the middle of the most highly conserved part of fusellovirus genomes (Figure 2-1). VP2 has DNA-binding activity (Reiter et al., 1987a; Iverson and Stedman, unpublished) that is presumably required for DNA packaging. ORF *b129* (NP_039795.1) was chosen because it is intolerant of insertional mutagenesis (Stedman et al., 1999), a high resolution structure is known (Lawrence et al., 2009) and the gene is completely conserved in all SSVs (Figure 2-1). Finally, ORF *d244* (NP_039781.1) was chosen for gene disruption because a high-resolution structure of its homolog from SSVRH is known (Menon et al., 2008) and it is conserved in most SSV genomes with the exception of SSVK1.

Materials and methods

Culture conditions

Sulfolobus solfataricus strains, Table 1, were grown aerobically at 76°C on plates or in liquid media containing yeast extract and sucrose as carbon and energy sources (YS Media), both as in Jonuscheit et al. (2003). *Escherichia coli* strains were grown in LB medium at 37°C as suggested by the manufacturer (Novagen).

Purification of DNA

Plasmid DNA used for LIPCR was purified from *E. coli* using the alkaline lysis method of Birnboim and Doly (1979). Plasmid DNA used to transform *Sulfolobus*

was purified using the GeneJet Plasmid Purification Kit (Fermentas) following the manufacturer's protocols. Total genomic DNA was isolated from *S. solfataricus* in late log phase growth (OD_{600nm} ~0.6) as in Stedman et al. (1999). Plasmid DNA was purified from a 50 mL culture of *S. solfataricus* transformed with SSV-Δ*d244* (late log, OD_{600nm} ~0.6) using the GeneJet plasmid purification kit (Fermentas) following the manufacturer's protocols. This DNA was retransformed into *E. coli* (Novagen), purified therefrom and analyzed by restriction endonuclease digestion with EcoRI (Fermentas).

Construction of SSV1 deletion mutants

Deletion mutants were constructed from the pAJC97 shuttle vector using LIPCR (Clore and Stedman, 2006). Primers were designed to overlap with the start and stop codon of the ORF to be deleted to keep the deletion in frame. Initially primers were designed using the Archaea genome browser (archaea.ucsc.edu). Primer melting temperatures were matched and then checked for potential primer dimer and secondary structure formation using online tools from IDT (Integrated DNA Technologies). Table 2 contains a list of oligonucleotide sequences used. LIPCR was performed using Phusion High-Fidelity DNA Polymerase (NEB/Finnzymes) at a final concentration of 0.005 U/μL. LIPCR cycling conditions as follows: initial denaturation at 98°C for 3 min; 35 cycles of 98°C for 15 s, annealing for 15 s, 72°C for 6 min, and a final extension at 72°C for 6 min. The annealing temperatures for deletion of *vp2*, ORF

d244, and ORF *b129* were 59, 53, and 66˚C, respectively. DNA was precipitated directly from LIPCR reactions using sodium acetate at a final concentration of 0.3 M and 95% EtOH. This DNA was phosphorylated using T4 polynucleotide kinase according to the manufacturer's protocols (Fermentas). DNA was ligated overnight (~20 h) at 16˚C using 5 Weiss units of T4 DNA ligase (Fermentas). Ligated DNA was transformed into NovaBlue Singles chemically competent *E. coli* following the manufacturer's protocol (Novagen). Plasmids were purified from single colonies and deletion constructs were identified by restriction endonuclease digestions. The deletion borders were confirmed by sequencing of the plasmids.

Electroporation of Sulfolobus

Purified plasmid DNA was electroporated into *Sulfolobus* strain G-theta as in Schleper et al. (1992). Following electroporation (400 Ω, 1.5 kV, 25 *μ*F) cells were immediately resuspended in 1 mL of YS media at 75º C and incubated for 1 h at º C. The cells were then added to 50 mL of pre-warmed YS media (75º C) and grown in liquid media as outlined below.

Screen for functional infectious virus/halo assay

To confirm the presence of infectious virus, halo assays were performed in duplicate 48 and 72 h post-electroporation (Stedman et al., 2003). Uninfected *Sulfolobus* G-theta cells were diluted to an OD_{600nm} = ~0.3 and allowed to grow until the OD_{600nm} reached ~0.35 (about 2.5 h). Half of a milliliter of this uninfected

culture was added to 5 mL YS media containing 0.2% wt/vol Gelrite as a softlayer and poured onto pre-warmed YS plates. Two microliters of supernatant from electroporated cultures was spotted onto the lawns and plates were incubated at 75°C for up to 3 days. A halo of host growth inhibition, typically observed 48– 72 h after incubation, indicated the presence of an infectious virus (Figure 2-2).

Growth curves

Portions of halos of growth inhibition from infected *S. solfataricus* G-theta cells were removed from plates with a sterile pipette tip and inoculated into liquid YS media. The culture was grown to an OD_{600nm} of ~0.6. One milliliter of this culture was diluted in 100 mL YS media to an OD_{600nm} ~0.050. Cultures were placed in a shaking incubator at 75°C and the OD_{600nm} was measured every 24 h. After 96 h, 1 mL of culture was diluted into 100 mL fresh YS media and returned to 75°C. One milliliter of culture was removed 72 h after each dilution, cells removed by centrifugation (14,000 rpm for 5 min in a microcentrifuge) and the supernatant was screened for virus using the halo assay above.

Transmission electron microscopy

Supernatant from infected cultures was collected by centrifugation at 14,000 rpm for 5 min in a microcentrifuge. Five microliters of supernatant were absorbed onto a 400 mesh carbon/formvar grid (Ted Pella) for 2 min and

negatively stained with 2% uranyl acetate for 20 sec. Grids were viewed on a JEOL 100CX TEM operated at 100 keV and images captured with a Gatan imager.

Results

SSV1 is infectious without the vp2 *gene*

The VP2 protein was purified from SSV1 virus particles and reported to be a DNA-binding protein (Reiter et al., 1987a). Surprisingly, a gene for VP2 was not found in SSV2 (Stedman et al., 2003), SSVRH or SSVK1 (Wiedenheft et al., 2004). Moreover, a homolog is not present in the *S. solfataricus* or *S. islandicus* genomes (She et al., 2001; Reno et al., 2009; Guo et al., 2011). However, a very distant relative of SSV1, SSV6, which also contains an atypical putative tail fiber protein, has a *vp2* gene (Redder et al., 2009). Thus, it is not clear whether SSV1 can function without a *vp2* gene.

Table 2-1: Strains and plasmid vectors used in this work

Strain/vector	Description	Reference
S. solfataricus GΘ	MT4 Derivative	Cannio et al. (1998)
S. solfataricus S443	Novel *Sulfolobus* isolate	Unpublished data
E. coli NovaBlue®	Expression strain	Novagen, Inc.
pAJC97	SSV1 with TOPO PCR Blunt II	Clore and Stedman (2006)
pAJC97-ΔVP2	pAJC97 lacking VP2 gene	This Work
pAJC97-Δd244	pAJC97 lacking ORF d244	This Work
pAJC97-Δb129	pAJC97 lacking ORF b129	This Work

Therefore, we made an in-frame deletion of the majority of the *vp2* gene by LIPCR in the context of the pAJC97 SSV1 shuttle vector (Clore and Stedman 2006), leaving the first four codons and the last four codons (including the stop codon) of the ORF intact (see Table 2-1). The putative promoter for the T9 "early" transcript was also left intact. The construct containing the deletion, pAJC97-Δ*vp2*, is hereafter referred to as SSV-Δ*vp2*.

To determine if the SSV-Δ*vp2* was able to make infectious virus, the shuttle vector was electroporated into *S. solfataricus* strain G-theta. Two days after electroporation, the supernatant from the transformed strains caused inhibition of growth of uninfected *S. solfataricus* strain G-theta on plates (Figure 2-2) that was indistinguishable from growth inhibition caused by the virus containing the *vp2* gene. Similar growth inhibition was also observed on lawns of uninfected *S. solfataricus* strain S443, a new *S. solfataricus* isolate from Lassen Volcanic National Park that is a host for all tested SSVs (Ceballos et al., 2012). Moreover, the supernatant contained SSV-like particles when observed by transmission electron microscopy (Figure 2-3).

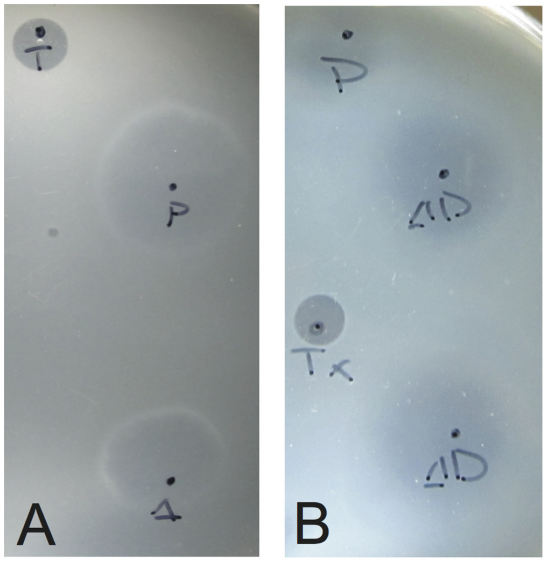


Figure 2-2: Typical growth inhibition of *S. solfataricus* on plates due to infectious virus
Lawns of *S. solfataricus* strain G-theta were prepared as in Stedman et al. (2003). Two microliters of supernatant from cultures transformed with either (A) SSV-Δ*VP2* or (B) SSV-Δ*d244* were placed on the lawns where indicated. Δ indicates where SSV-Δ*vp2* was spotted, ΔD where SSV-Δ*d244* was spotted. P indicates SSV-wild type spotted as a positive control. T or Tx indicates 2 µL of 0.01% Triton X-100 spotted as a control for lawn growth.

Infection by wild-type SSV1 and shuttle vectors does not drastically slow growth of cells in liquid culture for unknown reasons (Martin et al., 1984; Schleper et al., 1992; Stedman et al., 1999). The same is true of SSV-Δ*vp2* (Figure 2-4). Infection with SSV-Δ*vp2* was confirmed via PCR amplification (data not shown).

SSV1 constructs lacking the conserved ORF b129 *do not appear to make infectious viruses*

The *b129* ORF in SSV1 is universally conserved in all fuselloviruses (Redder et al., 2009). Moreover, shuttle vectors with pBluescript inserted into ORF *b129* did not produce infectious virus when electroporated into *Sulfolobus* (Stedman et al., 1999). However, a similar insertion mutant in the equally

conserved SSV1 viral integrase appears to be non-functional (Stedman et al., 1999), but an in-frame deletion was functional (Clore and Stedman, 2006). A structure for the *b129* ORF is also known (Lawrence et al., 2009) and it contains two Zn-finger putative DNA-binding motifs.

Table 2-2: Oligonucleotides used in this work

Name	Sequence	Description
VP2 LIPCR F	5′-CAC CGC AAG TAG GCC-3′	Flanks VP2 gene for deletion
VP2 LIPCR R	5′-CAC CCA CTT CAT ATC ACT CC-3′	Flanks VP2 gene for deletion
d244 LIPCR F	5′-ATC CAT TTA CCA TAA TCC ACC-3′	Flanks ORF d244 for deletion
d244 LIPCR R	5′-GGA AAA TGA TAT TCA ACT CAG AGG-3′	Flanks ORF d244 for deletion
b129 LIPCR F	5′-AGT TAG GCT CTT TTT AAA GTC TAC C-3′	Flanks ORF b129 for deletion
b129 LIPCR R	5′-TGA CTC CGT CAT CCT CTA AC-3′	Flanks ORF b129 for deletion
VP2 Check F	5′-ATT CAG ATT CTG WAT WCA GAA C-3′	Amplifies VP2 gene and flanking sequences
VP2 Check R	5′-TCS CCT AAC GCA CTC ATC-3′	Amplifies VP2 gene and flanking sequences
d244 Check F	5′-GGA ACT CCT CTC ATT AAC C-3′	Amplifies ORF d244 and flanking sequences
d244 Check R	5′-GAT CAT CAA CGA GTA TAT TGA CC-3′	Amplifies ORF d244 and flanking sequences
b129 Check F	5′-ATG AAG GCT GAG GAA ACA ATC GTG-3′	Amplifies ORF b129 and flanking sequences
b129 Check R	5′-TTA ATA TAG CTG CGA TGC AGT ATA GTT TAT TTG TGC-3′	Amplifies ORF b129 and flanking sequences

**Underlined sequence indicates ORF.*

The *b129* ORF was deleted with LIPCR leaving the first four and last two codons of the ORF intact and maintaining the predicted T3 promoter (Reiter et al., 1987b). This construct is referred to as SSV-Δ*b129*. Unlike the SSV-Δ*vp2* construct, supernatants from *Sulfolobus* cells electroporated with SSV1-Δ*b129* did not cause zones of growth inhibition when spot- ted on lawns of uninfected *S. solfataricus* strain G-theta. A total of nine independent transformations were performed in which the wild-type virus consistently caused growth inhibition but SSV- Δ*b129* did not. Moreover, no halos of growth inhibition were formed on lawns of *S. solfataricus* strain S443. It is not currently known at which step of virus replication the SSV-Δ*b129* is deficient.

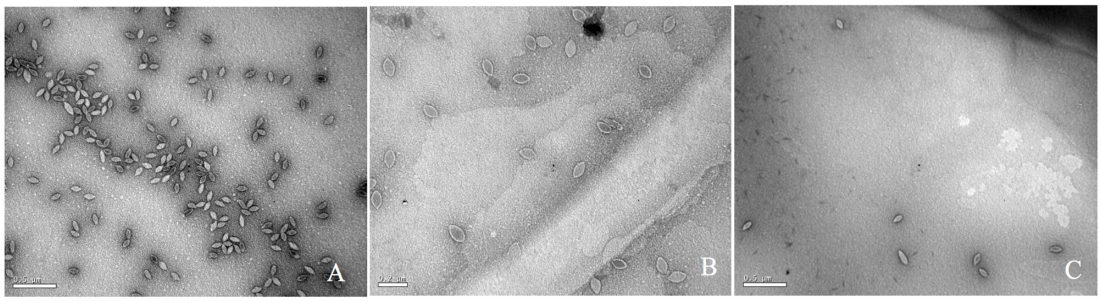


Figure 2-3: Transmission electron micrographs of SSV particles
Supernatants from cultures of *S. solfataricus* strain G-theta transformed with (A) pAJC97, (B) SSV-Δ*vp2*, (C) SSV-Δ*d244*, were negatively stained with uranyl acetate and observed with a JEOL 100CX transmission electron microscope. Bar represents 0.2 µm (B) or 0.5 µm (A,C).

SSV1 lacking ORF d244 *is infectious but has a novel phenotype*

SSV1 ORF *d244* is encoded on the transcript T5, upstream of the viral integrase gene (Figure 2-1). The entire pBluescript plasmid can be inserted into the ORF directly upstream of ORF *d244* without abrogating SSV1 function (Stedman et al., 1999). ORF *d244* is well conserved in other fusellovirus genomes with the exception of SSVK1 (Wiedenheft et al., 2004; Redder et al., 2009). The X-ray crystal structure of the homolog of SSV1 ORF *d244*, SSVRH ORF *d212* has been solved and it is predicted to be a nuclease (Menon et al., 2010). Moreover, the product of ORF *d244* has been reported to be in purified SSV1 particles (Menon et al., 2008).

The SSV1 *d244* ORF was deleted with LIPCR. The deletion of the *d244* ORF left the first two and last three codons of the ORF intact as well as maintained the ORF to avoid polar effects. This construct is referred to as SSV-Δ*d244*.

To determine if SSV-Δ*d244* was able to make infectious virus, the shuttle

vector was electroporated into *S. solfataricus* strain G-theta. Two days after electroporation, the supernatant from the transformed strains caused inhibition of growth of uninfected *S. solfataricus* strain G-theta on plates (Figure 2-2) and also inhibited growth of *S. solfataricus* strain S443 (data not shown). The supernatant contained SSV-like particles when observed by transmission electron microscopy (Figure 2-3).

Infection by wild-type SSV1, shuttle vectors and SSV-ΔVP2 does not slow growth of cells in liquid culture (Martin et al., 1984; Schleper et al., 1992; Stedman et al., 1999; see above). However, infection by SSV-Δ*d244* drastically slows growth of *S. solfataricus* strains G-theta and S443 in liquid culture (Figure 2-4). Infection with SSV-Δ*d244* was confirmed via PCR. Moreover, restriction endonuclease digestion of viral DNA recovered from transformed *S. solfataricus* cells and retransformed into *E. coli* revealed no obvious alterations of the SSV-Δ*d244* construct (data not shown).

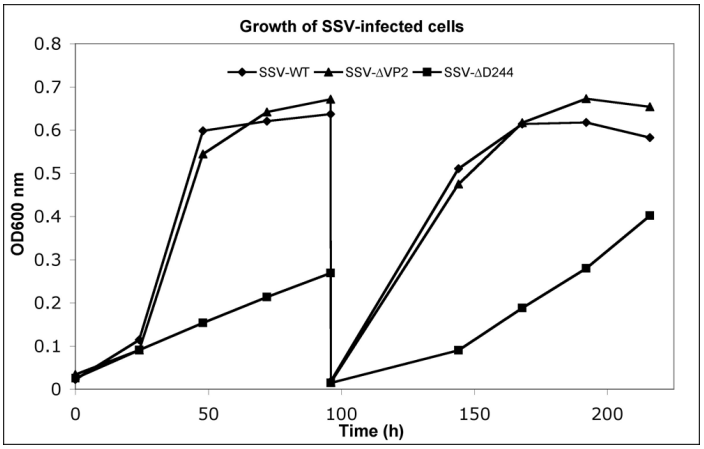


Figure 2-4: Typical growth inhibition in liquid culture of virus constructs
Cultures of *S. solfataricus* G-theta infected with wild-type SSV1, diamonds, SSV-Δ*vp2*, triangles and SSV-Δ*d244*, squares, were diluted in YS media to equal starting OD_{600nm} and incubated at 75°C. At the indicated times, samples were removed and the OD_{600nm} was determined and the presence of virus was confirmed in each culture via halo assay. After 96 h, 1 mL of cells were diluted 1:100 in fresh YS media and returned to 75°C.

Discussion

The putative DNA packaging protein VP2 is not required for SSV1 function

The deletion of *vp2* from SSV1 results in a functional virus that is indistinguishable from the wild-type virus (Figures 2-2 – 2-4). Based on the lack of conservation of VP2 this result is not completely unexpected. However, almost all viruses contain a genome packaging protein. There is no clear sequence homolog of VP2 in the host genome, but there are a number of small DNA-binding proteins, such as Sso7d or Cren7 that may be able to functionally substitute for VP2 in SSV1 genome packaging (Choli et al., 1988; Guo et al., 2008). This will be tested with mass spectrometry of SSV- Δ*vp2* particles. Alternatively, the VP2 protein may be involved in maintenance of the positive supercoiling of the SSV1 viral genome (Nadal et al., 1986). It would be interesting to know if the topology of the viral DNA is affected by the absence of VP2. It is predicted that positive supercoiling should increase the thermal stability of the DNA, so SSV-Δ*vp2* may be less thermally stable than the wild-type virus.

The *vp2* gene may be more prevalent than previously thought. *vp2*-like sequences have been reported from metagenomic studies, one in an acid mine drainage metagenome (Andersson and Banfield, 2008) and the other from Boiling Springs Lake in California (Diemer and Stedman, unpublished). These

vp2 genes may be in the context of a SSV6 or ASV-like genome (Redder et al., 2009).

The product of ORF b129 *appears to be essential for SSV1 infectivity*

Homologs of SSV1 ORF *b129* are present in all known SSVs (Redder et al., 2009). The *b129* ORF also does not tolerate insertion of the pBluescript plasmid (Stedman et al., 1999). Thus, it is not surprising that deletion of ORF *b129* leads to an incompletely replicating virus. However, the SSV1 integrase, a gene also conserved in all fuselloviruses, did not appear to tolerate insertion of pBluescript (Stedman et al., 1999), but could be deleted with LIPCR without abrogating virus function (Clore and Stedman, 2006). This indicates that either polar effects are important, which seems unlikely since the SSV1 *integrase* is at the end of the T5 transcript, or that insufficient replicate transformations were performed in the earlier study.

Nine replicate transformations of *S. solfataricus* with SSV- Δ*b129* did not generate functional virus. However, we cannot absolutely determine that SSV1 ORF *b129* is essential for virus function without complementation experiments, which are underway. The reasons for the apparent necessity of SSV1 ORF *b129* are unclear, but the structure of the *b129* ORF product, a predicted transcriptional regulator (Lawrence et al., 2009) and induction of the T6 transcript containing ORF *b129* after UV- irradiation (Reiter et al., 1987b; Fröls et al., 2007) provides clues to its function.

The assay used herein for virus infection, ability to cause a zone of growth inhibition on a lawn of uninfected cells, is for virus spread and infectivity. There are many other aspects of virus replication that could be affected by disruption of ORF *b129*. An attractive hypothesis is that the B129 protein activates transcription of virus structural genes encoded by the "late" transcripts T7/8/9, T1, and T2 (Reiter et al., 1987b; Fröls et al., 2007; see Figure 2-1). This would be one of very few archaeal transcriptional activators characterized to date and the only the second archaeal viral transcriptional activator (Kessler et al., 2006). Thus, the SSV-Δ*b129* construct may be able to replicate its genome, integrate into the host, and have genome replication induced by UV-irradiation or some subset of these activities. Experiments to test these hypotheses are underway.

*Transfection with SSV-Δ*d244 *produces virus and retards cell growth*

The SSV1 *d244* ORF is well-conserved in fuselloviruses with the exception of SSVK1 (Wiedenheft et al., 2004; Redder et al., 2009). However, SSV1 lacking ORF *d244* clearly makes infectious virus particles (Figures 2-2 and 2-3). Moreover, the zones of clearing produced by supernatants of cells transfected with SSV-Δ*d244* are clearer than those produced by either the wild-type or SSV-Δ*vp2* viruses (Figure 2-3; unpublished data). They are reminiscent of zones of clearing produced by SSVK1 (data not shown). Unlike wild-type virus and SSV-Δ*vp2*, transfection by SSV-Δ*d244* leads to drastically reduced host growth (Figure 2-4). The reasons for this growth inhibition are unclear. Similar

growth phenotypes have been observed in SSVK1 infections (Stedman et al., in preparation). SSVK1 consistently produces more virus than similar cultures of the wild-type virus, so this may account for the growth defect (unpublished data). Whether SSV-Δ*d244* consistently produces more virus than the wild-type or SSV-Δ*vp2* is currently unknown.

The structure of the product of SSV1 ORF *d244* is a predicted nuclease (Menon et al., 2010), similar to Holiday junction resolvase enzymes. Why the lack of a resolvase leads to slower host growth is unclear. Possibly SSV1 ORF *d244* is involved in the specificity of SSV1 integration. SSV-K1 is known to integrate into multiple positions in the host genome (Wiedenheft et al., 2004), which may contribute to its higher copy number. Whether SSV-Δ*d244* integrates into multiple positions in the host genome is under investigation. On the other hand, there may be a defect in SSV-Δ*d244* replication or resolution of SSV replication intermediates that leads to accumulation of aberrant DNA, which, in turn, leads to slower host growth.

After multiple transfers of *Sulfolobus* cultures transfected with SSV-Δ*d244* into fresh media, growth rates recover to near wild- type rates (unpublished data). The virus is still present in these cultures by PCR and is able to inhibit *Sulfolobus* growth on plates (unpublished data) so the virus is not lost or apparently rearranged (see Results). Whether there are other genetic changes in the virus or host under these conditions remains to be determined. One attractive possibility is changes to the CRISPR repeat structures that are proposed to be

important for acquired immunity in *Sulfolobus* (Held and Whitaker, 2009).

Summary and outlook

Comparative and structural genomics has identified a number of targets for gene disruption in the SSV1 genome. Here precise gene disruptions of the poorly conserved *vp2* gene, and the well-conserved ORFs *b129* and *d244* are described. Deletions in *vp2* may allow insights into DNA packaging in the SSV1 genome. Deletion of ORF *b129* may allow the identification of the second archaeal virus transcriptional activator. Deletion of ORF *d244* may allow insight into copy number regulation in SSVs, previously thought to be regulated by ORF *d63* (Lawrence et al., 2009). Clearly, there are many more genes to be analyzed in the SSV1 genome and more insights that can be gained by combining comparative genomics, structural biology, and genetics. In the future, biochemical work will be added to this suite of techniques to gain fundamental understanding of this fascinating, unique, and ubiquitous archaeal virus family.

Chapter 3: A comprehensive genetic survey of the fusellovirus SSV1 reveals a malleable genome that tolerates deletion of the minor capsid gene

Abstract

The viruses infecting the Archaea harbor a tremendous amount of genetic diversity. This is especially true for the spindle-shaped viruses from the family *Fuselloviridae*, where >90% of the viral genes do not yield detectable homologs in public databases. This significantly limits our ability to elucidate the role of viral proteins in the infection cycle. To address this, we have developed genetic techniques to study the well-characterized fusellovirus *Sulfolobus* spindle-shaped virus 1 (SSV1).

Here, we expand upon previous research and present a thorough genetic analysis of SSV1 using specific and random mutagenesis. We demonstrate that almost half of the SSV1 genes are not essential for infectivity and the requirement for a particular gene roughly correlates with its degree of conservation within the *Fuselloviridae*. Surprisingly, the universally conserved minor capsid gene *vp3* could be deleted without a loss in infectivity and results in slightly elongated virions. The major capsid gene *vp1* is essential for SSV1 infectivity but does tolerate mutations to a well-conserved glutamate believed to be important for proteolytic processing of the protein. SSV1 *vp1* deletion mutants could be rescued by complementation in cis with *vp1* homologues from SSV2 and SSV9, a technique that has not been demonstrated for an archaeal virus.

Introduction

The Archaea are infected by some of the most structurally and genetically diverse viruses known in nature (Reviewed by Prangishvili 2013). Of the numerous morphologically unique viruses, those possessing a spindle-shaped architecture are widespread within the Archaea and include some of the best characterized archaeal viruses. Spindle-shaped viruses (SSVs) belong to two viral families, the *Fuselloviridae* and *Bicaudoviridae,* which share few similarities other than overall capsid morphology (Krupovic et al., 2014). The *Fuselloviridae* are further divided into the genera *Alphafuselloviridae* (SSV1, SSV2, SSV4, SSV5, SSV7, SSV8, and SSV9) and the *Betafuselloviridae* (SSV6 and ASV1) (ICTV 2014). Fuselloviruses have been isolated world-wide from volcanic hot spring environments (70-80^{o}C, pH~2-4) in which their hosts thrive (Yeats et al., 1982; Stedman et al., 2003; Redder 2009). All members of the *Fuselloviridae* possess genomes of similar size and content, however variations in size and morphology of the particle exist (Redder et al., 2009). *Sulfolobus* spindle-shaped virus 1 (SSV1) is the best-studied fusellovirus, although a significant amount of research has been done on the closely-related SSV2 (Reviewed in Contursi et al., 2014).

SSV1 and its host *S. shibatae* were originally isolated from a hot spring in Beppu, Japan (Yeats et al., 1982). SSV1 possesses a circular positively supercoiled 15.5 kbp dsDNA genome (Palm et al., 1991; Nadal et al., 1986). SSV1 is apparently unique among the *Fuselloviridae* in that its replication is

strongly UV inducible (Martin et al., 1984; Fusco et al., 2015a). Transcription following UV induction was analyzed and found to be temporally regulated (Reiter et al., 1987b; Fröls et al., 2007). All fuselloviruses encode a tyrosine family recombinase that facilitates integration into a host tRNA gene, although integration does not seem to be required for virus infectivity (Clore and Stedman 2006). Because SSV1 integrates into the host genome and establishes a stable carrier-state, SSV1 has been referred to as a temperate virus (Fusco et al., 2013). This nomenclature can be misleading, however, as SSV1 maintains a low copy number of episomal DNA (~5 copies/cell) and SSV1 infections do not appear to result in lysis of the host cell (Fusco et al., 2013; personal observations). SSV1 virions appear to be released via budding and not through the recently discovered novel viral egress mechanism utilized by SIRV2 and STIV1 viruses (Martin et al., 1984; Bize et al., 2009; Brumfield et al., 2009).

Fusellovirus capsids are predominantly comprised of the major capsid protein VP1 and small amounts of the minor capsid protein VP3 (Reiter et al., 1987a; Quemin et al., 2015). VP1 and VP3 proteins are highly conserved within the *Fuselloviridae* and share significant homology to each other (Reiter et al., 1987a; Iverson and Stedman 2012). VP1 isolated from purified SSV1 virions indicates that VP1 is proteolytically cleaved at an internal universally conserved glutamate prior to assembly (Reiter et al., 1987a; Quemin et al., 2015). The source and identity of this putative protease remains a mystery. SSV1 and three other fuselloviruses also encode a small and extremely basic structural protein

(VP2) that is thought to bind to viral DNA within the capsid (Reiter et al., 1987a; Servin-Garcidueñas et al., 2013). A fourth capsid protein, VP4 (formerly C792), has been identified via mass spectrometry of purified virions and likely forms the tail fiber of the virion (Menon et al., 2008; Quemin et al., 2015). Recently it was shown that purified SSV1 virions contain host-derived glycerol dibiphytanyl glycerol tetraether (GDGT) lipids (Quemin et al., 2015). The structure of the SSV1 virion has been solved via cryoelectron microscopy, providing insight into the architecture and assembly of this diverse family of viruses (Stedman et al., 2015; Quemin et al., 2015).

The genome of SSV1 contains 35 open-reading frames (ORFs), most of which cannot be assigned a function due to undetectable homology with sequences in public databases (Iverson and Stedman 2012; Chapter 2). This situation is not unique among viruses, however it seems especially prolific among crenarchaeal viruses and hinders our understanding of them and their life cycles (Prangishvili et al., 2006). Comparison of fusellovirus genomes have identified a set of 12 genes/ORFs (i.e. the fusellovirus core) that are completely conserved within the *Fuselloviridae* (Table 1-1). These genes are almost entirely clustered on one half of the genome and are predicted to play roles in replication, assembly and packaging due to the timing of their transcription in the UV-induced transcription cycle (Fröls et al., 2007; Redder et al., 2009). To date, only the SSV1 structural proteins (VP1, VP2, VP3, and VP4) and integrase have been assigned a function, although a combination of structural studies and

bioinformatics have provided predictions for the roles of several others (Figure 3-1).

A number of putative transcription factors have been identified in the SSV1 genome, none of which have been experimentally characterized. Crystal structures of SSV1-B129, SSV1-F112, SSV1-F93 and SSV8-E73 (homologue of SSV1-E51) reveal DNA binding domains characteristic of transcriptional regulators (Lawrence et al., 2009; Menon et al., 2008; Kraft et al., 2004a; Schlenker et al., 2012). The products of four other well conserved SSV1 ORFs (SSV1-A45, SSV1-A79, SSV1-C80, and SSV1-B115) are predicted to encode ribbon-helix-helix (RHH), helix-turn-helix (HTH), or zinc finger (ZNF) DNA binding domains and may also be involved in transcription regulation (Prangishvili et al., 2006). The protein product of ORF *f55* is predicted to possess a RHH DNA binding domain and was shown to weakly bind a number of viral promoters that control the expression of "early" gene products. Based on these data, SSV1-F55 is hypothesized to maintain "lysogeny" by repressing early viral promoters (Fusco et al., 2013; Fusco et al., 2015b), but this is inconclusive.

SSV1-B251 and SSV1-A153 are highly conserved and are the only fusellovirus proteins encoded by the satellite virus pSSVx, leading to a prediction they may be involved in replication or packaging (Arnold et al., 1999). In line with this, SSV1-B251 possesses clear NTP binding motifs and is predicted to be homologous to bacterial DnaA (Koonin et al., 1992). The structure of SSV8-D212, a homologue of SSV1-D244, displays a nuclease fold although activity

was not demonstrated biochemically (Menon et al., 2008). SSV1-D244 was identified by mass spectrometry on purified SSV1 virions, although it was not identified in a later analysis (Menon et al., 2008; Quemin et al., 2015). The structure of SSV1-D63 displays a four-helix bundle that is characteristic of a large number of proteins making a functional prediction difficult (Kraft et al., 2004b).

SSV1 has proven to be quite amenable to genetic manipulation and is one of only two crenarchaeal viruses with established genetic systems (Stedman et al., 1999; Wirth et al., 2011). The ability of SSV1 to tolerate large insertions of foreign DNA allowed for the construction of SSV1-based shuttle vectors and provided tools for studying the viral genome itself (Stedman et al., 1999; Jonuscheit et al., 2003). Digestion of SSV1 DNA followed by the insertion of an *E. coli* plasmid provided the first data of which genes are required for virus infectivity (Stedman et al., 1999). The data indicated that a majority of ORFs in SSV1 are apparently essential for virus function. To date SSV1 is the only crenarchaeal virus to be extensively studied via genetic manipulation, although SSV9 has been shown to tolerate insertion of foreign DNA (Stedman unpublished; Appendix A).

SSV1 shuttle vectors were essential for subsequent work where specific ORFs were deleted from the viral genome. Using long inverse PCR (LIPCR) the SSV1 *integrase* gene was deleted and subsequently shown to be non-essential for production of infectious virus in *S. solfataricus* P2 (Clore and Stedman 2006). As a follow up to this work, the *vp2* gene and ORFs *b129* and *d244* were deleted

(Iverson and Stedman 2012; Chapter 2). Deletion of *b129* resulted in loss of infectivity whereas removal of *vp2* and *d244* resulted in production of infectious virus. Homologues of SSV1-B129 have been identified in the genomes of all 11 isolated fuselloviruses, indicating that it is probably crucial for virus infectivity (Appendix B). SSV1-VP2 is much less well conserved and may explain why it can be deleted without a loss of infectivity (Appendix B). Interestingly, cells that were infected with Δ*d244* virus exhibited a much more retarded growth phenotype when compared to cells infected with wild-type or Δ*vp2* viruses (Iverson and Stedman 2012; Chapter 2).

In an effort to further characterize the genetic requirements for the function of SSV1, LIPCR and transposon mutagenesis were employed to construct individual SSV1 genomes harboring defects in each of the 35 ORFs. Overall, SSV1 appears to be much more tolerant of mutagenesis than previously thought. Almost the entirety of the T5 early transcript appears to be expendable while the other early transcript is much less so, in correlation with the abundance of well conserved fusellovirus genes. The genes of the fusellovirus core appear to be essential with the surprising exception of the minor capsid gene *vp3*, which was shown to be non-essential for SSV1 infectivity. Complementation of an SSV1 mutant is also demonstrated for the first time.

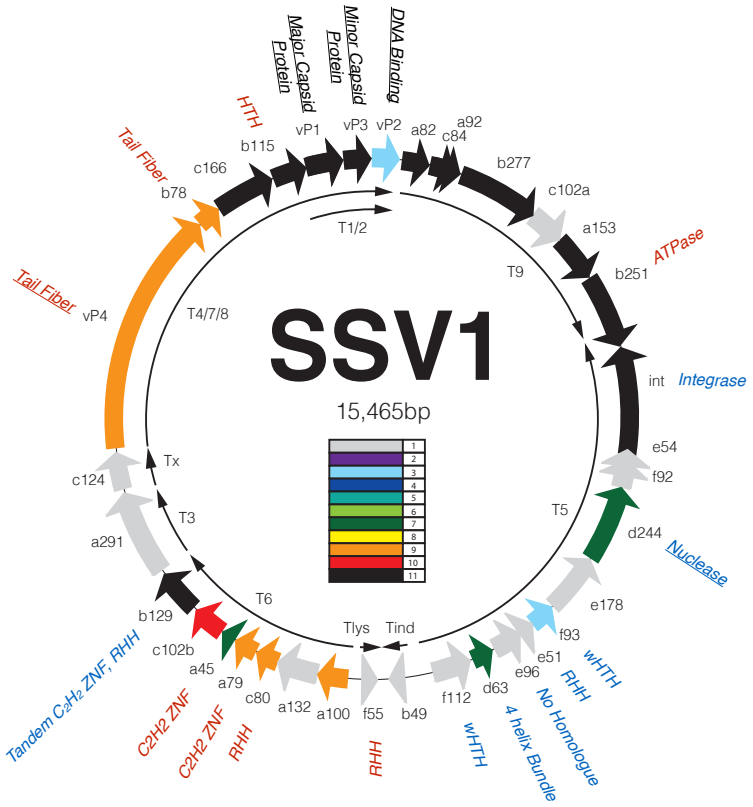


Figure 3-1: Genome map of SSV1
Open reading frames (ORFs)are displayed as block arrows and labeled as in Palm et al., 1991. The conservation of each ORF among 11 fusellovirus genomes is indicated by the color code in the middle of the map. Homologous ORFs were determined by pBLAST (NCBI) using an e-value cutoff of 0.001. Functional and structural annotations are listed next to the ORF name. Annotations in blue denote proteins that have been structurally characterized (Integrase: Eilers et al., 2012; D244: Menon et al., 2010; F93: Kraft et al., 2004a; E51: Schlenker et al., 2012; E96: Unpublished, PDB ID 2WBT; D63: Kraft et al., 2004; F112: Menon et al., 2008; B129: Lawrence et al., 2009). Annotations in red denote ORFs that were characterized via bioinformatics (*vp4* and *b78*: Redder et al., 2009; *a45*, *a79*, *c80* and *b115*: Prangishvili et al., 2006; b251: Koonin et al., 1992; *f55*: Fusco et al., 2013). Underlined annotations indicate proteins that have been identified within the purified SSV1 virion (Reiter et al., 1987a; Menon et al., 2008; Quemin et al., 2015). Viral transcripts are shown as small arrows on interior of map (Reiter et al., 1987b; Fröls et al., 2007; Fusco et al., 2013). HTH = helix-turn-helix; RHH = ribbon-helix-helix; wHTH = winged helix-turn-helix; ZNF = zinc finger.

Materials and Methods

Purification of SSV1 DNA from Sulfolobus*:*

S. shibatae strain B12, the original SSV1 isolate (Yeats et al., 1982; all microbial strains used are listed in Table 3-1, was grown by adding small "chip" of frozen cell stock (ca. 50 μL) to 5 mL of YS medium and incubated at 70º C with shaking until turbid (about ten days). This culture was transferred to 50 mL fresh YS and grown for 3 days (OD_{600nm} = 0.75) after which the culture was centrifuged at 3,000 x g. The virus containing supernatant was stored at 4º C and SSV1 DNA was purified from the pellet using the GeneJet Plasmid Purification kit following the manufacturer's protocol for plasmid purification from *E. coli* (Thermo-Fisher). The yield of purified plasmid was consistently low and had a large amount of background DNA contamination, presumably from degradation of the host chromosome. Purification using only the GeneJet kit yields DNA suitable for PCR, however, the total SSV1 DNA yield was low and deemed too impure for transposon mutagenesis. An SSV1 DNA sample of higher purity was obtained as follows. 15 mL of uninfected *Sulfolobus* strain S441 (Table 3) was pelleted, resuspended in 15 mL of SSV1 supernatant (from infected B12 cells), and incubated at room temperature with shaking for two hours. These infected cells were then added to 50 mL fresh YS medium and grown for 72 hours at 70º C. The culture was spun at 3,000 x g and the pellet was resuspended in 400 μL of resuspension buffer (50 mM Glucose, 25 mM Tris, pH 8.0) and split into four 100 μL aliquots. 200 μL of lysis buffer (0.2 M NaOH,

1% SDS) was added to each aliquot, mixed, and incubated on ice for 5 minutes. 150 μL of neutralization buffer (3M potassium acetate, 2M Acetic acid) was added and samples were incubated on ice for 5 minutes. Samples were spun for 10 min and the supernatant was collected and extracted three times with an equal volume of phenol:chloroform:isoamyl-alcohol (PCA).

The aqueous phases of PCA extractions were pooled and DNA was precipitated with ethanol. PCA contaminants were removed by passing DNA through a GeneJet Plasmid Purification column following the manufacturer's protocol. The resulting DNA was analyzed by UV spectroscopy (260/280 nm ~1.8) and EcoRI endonuclease digestion according to manufacturer's protocol (Thermo-Fisher).

Table 3-1: *Sulfolobus* and *E. coli* strains used in this work

Strain	Description	Reference
S. shibatae B12	Original SSV1 host	Yeats et al., 1982
S. solfataricus S441	Novel *Sulfolobus* isolate	Unpublished
S. solfataricus S171	Original SSV9 host	Wiedenheft et al., 2004
S. solfataricus GΘ	MT4 Derivative	Cannio et al., 1998
Transformax EC100D pir$^+$ *E. coli* [A]	*F⁻ mcrA Δ(mrr-hsdRMS-mcrBC) ϕ80dlacZΔM15 ΔlacX74 recA1 endA1 araD139 Δ(ara, leu)7697 galU galK λ- rpsL (Str^R) nupG pir$^+$(DHFR)*	Epicentre, Inc
Transformax EC100D pir-116 *E. coli* [A]	*F⁻ mcrA Δ(mrr-hsdRMS-mcrBC) ϕ80dlacZΔM15 ΔlacX74 recA1 endA1 araD139 Δ(ara, leu)7697 galU galK λ- rpsL (Str^R) nupG pir-116(DHFR)*	Epicentre, Inc
NovaBlue Singles *E. coli* [A]	*endA1 hsdR17* (rK12– mK12+) *supE44 thi-1 recA1 gyrA96 relA1 lac* F′[*proA+B+ lacIqZΔM15*::Tn*10*] (TetR)	Millipore, Inc

[A] Chemically competent cells were prepared in house (See methods)

Purification of LIPCR template DNA from E. coli

E. coli cells harboring the pAJC97, EAI283, or EAI228 SSV1 shuttle vectors (see Table 3-3 for reference to plasmids and vectors) were grown overnight in 5 mL of LB medium containing 50 μg/mL Kanamycin. Plasmid DNA was then purified from 1.5 mL of overnight culture via alkaline lysis (see above for buffer components; Birnboim and Doly 1979). Following ethanol precipitation, the DNA was dissolved in 30 μL diH_2O with 0.01 μg of RNaseA. DNA was analyzed by EcoRI restriction endonuclease digestion following manufacturer's protocol (Thermo-Fisher). DNA was typically diluted 1:10, 1:50, and 1:100 in diH_2O for use as template in LIPCR.

Long inverse PCR (LIPCR)

LIPCR was used to delete entire SSV1 ORFs, portions of ORFs, or to introduce single base pair mutations into the SSV1 genome. For deletions, primers were designed to overlap with the start and stop codons of the ORF to be deleted. Due to complications with primer design (incompatible Tm's, unfavorable secondary structures, primer dimers), most primers include additional portions of the 5'/3' ends of ORFs (Table 3-2). For *vp1* single base pair mutations, primers were designed in such a way that ligation of the LIPCR product resulted only in desired base pair mutation and no deletion. Template DNA for LIPCR was purified via alkaline lysis from *E. coli* harboring the pAJC97, EAI283, or EAI228 SSV1 shuttle vectors (described above; Table 3-3). The

optimal concentration of template DNA for LIPCR was unknown and had to be determined empirically for each set of primers. Template DNA purified from *E. coli* was initially diluted in 30 μL of H_2O with 0.01 μg RNaseA and further diluted 1:10, 1:50, 1:100 in H_2O. One μL of each dilution was used as template in a 20 μL reaction and the dilution resulting in the cleanest full length product was selected. LIPCR was performed as described in Iverson and Stedman 2012 (Chapter 2). Annealing temperatures for each primer pair were estimated using NEB Tm prediction software (http://tmcalculator.neb.com/#!/) but typically had to be experimentally optimized. LIPCR products were purified, ligated, and transformed into chemically competent NovaBlue *E. coli* as in Iverson and Stedman 2012 (Chapter 2).

Transposon mutagenesis

The EZ-Tn5^{TM} <R6Kγori/KAN-2> Insertion Kit (Epicentre) was used to perform transposon mutagenesis on purified SSV1 DNA. A molar ratio of 30:1 SSV1 DNA to EZ-Tn5 transposon was found to yield significantly more mutants than the manufacturer's recommended equimolar ratio. This was the only deviation from the manufacturer's protocol (Epicentre). Following completion of the reaction, the sample was diluted to 20 μL with nuclease free H_2O. This was done to lower the conductivity of the sample and reduce the chance for arcing during electroporation. 1 μL of the SSV1-EZ-Tn5 reaction was electroporated into 50 μL of transformax EC100D pir^{+} electrocompetent *E. coli* (Epicentre; Table

3-1) using a 0.1 cm gap length cuvette (Bulldog Bio) under the following electroporator (BioRad Gene Pulser II) conditions: 1.8 kV, 200 Ω, 25 μF. Time constants were generally around 3.5-4.0 milliseconds. Electroporated cells were immediately transferred to 1 mL of SOC medium and allowed to recover for 1 hr at 37º C with shaking. Cells were then spread on LB plates containing 50 μg/mL Kanamycin and grown overnight at 37º C.

Preparation of chemically competent E. coli for transformation of SSV1 deletion mutants

NovaBlue Singles *E. coli* (Millipore; Table 3-1) were the only cells capable of consistent transformation with the large plasmid pAJC97, from which almost all SSV1 deletion mutants were constructed. To prepare chemically competent cells, a single colony of NovaBlue *E. coli* from a freshly streaked plate was inoculated into 4 mL of sterile LB medium with no antibiotic and grown overnight at 37º C with shaking. The culture was transferred to 400 mL sterile LB and typically grown to an OD_{600nm} ~0.3 - 0.4. Cells were spun for 10 min and resuspended in 1/4 volume of ice cold 100 mM $MgCl_2$. Cells were then spun for 10 min, resuspended in 1/2 vol of ice cold 100 mM $CaCl_2$, and left on ice for 30 min. Cells were centrifuged for 10 min and resuspended in 85 mM ice cold $CaCl_2$/15% Glycerol. Cells were divided into 100 μL aliquots and frozen at -80ºC. Cells retained competence for greater than 6 months at -80º C. All centrifugations were performed at 3,000 x g and 4º C. Chemically competent

Transformax EC100D pir$^+$ cells used for transformations of EZ-Tn5 based vectors were prepared in an identical fashion (Table 3-1).

Isolation and identification of transposon and deletion mutants

Single colonies of *E. coli* following transformation with LIPCR products or transposon insertion reactions were picked and grown in 5 mL LB with 50 μg/mL Kanamycin. Plasmid DNA was purified via alkaline lysis (see above) and checked for a SSV1-like banding pattern after digestion with EcoRI (Thermo-Fisher). Deletion mutants displaying expected EcoRI restriction patterns were further analyzed by PCR using primers that amplified across the deleted region. Positively identified mutant DNA was re-isolated from *E. coli* using the GeneJet Plasmid Purification Kit (Thermo-Fisher) following manufacturer's protocol. Purified DNA was sequenced using the Big Dye Terminator Kit according to manufacturer's protocols (ABI). 10 μL BDT sequencing reactions were assembled as follows: 200-400 ng template DNA, 0.5 μL of 10 μM primer, 2 μL 5X BDT buffer, 1 μL of BDT mix, and H_2O to 10 μL. Thermocycler parameters for BDT sequencing reactions were as follows: Initial denaturation at 96º C for 2 min; 30 cycles of 10 sec at 96º C, 10 sec annealing at 50º C, 4 min extension at 72º C. BDT reactions were precipitated with isopropanol and allowed to dry at room temperature. Samples were submitted to the DNA Services Core at Oregon Health and Sciences University for sequencing.

Complementation in cis of SSV1 mutants

SSV1 mutants lacking a specific ORF/gene were constructed via LIPCR using EAI283 as template (Tables 3-3 and 3-4). The homolog of the missing ORF/gene from either SSV9 or SSV2 (Table 3-4) was amplified using Phusion DNA polymerase, purified with PCR Purification Kit (Thermo-Fisher) and phosphorylated with T4 polynucleotide kinase according to manufacturer's protocols (Thermo-Fisher). Phosphorylated homolog-containing DNA was heat treated for 10 min at 75º C prior to ligation reaction. LIPCR products were purified directly from the LIPCR reaction by sodium acetate/ethanol precipitation. The phosphorylated homolog-encoding DNA was mixed with LIPCR product at a molar ratio of 10:1 (respectively) and ligated with 5 U of T4 DNA ligase (Thermo-Fisher) for 20 hours at 16ºC. 5 μL of the ligation reaction was used to transform 100 μL chemically competent pir$^+$ *E. coli* according to manufacturer's protocol (Epicentre). DNA was purified from resulting transformants by alkaline lysis and checked for correct insertion by PCR with DreamTaq DNA polymerase (Thermo-Fisher) using an internal primer specific to the homolog and an external primer specific to SSV1. BDT sequencing was performed to confirm correct insertion (see above).

Transformation of Sulfolobus

Starter cultures of *Sulfolobus* (5 mL) were grown from frozen stocks for 48-72 hr in a 70º C shaking incubator. Starter cultures were transferred to 50-

100 mL of fresh YS in long neck Erlenmeyer flasks and grown until turbid (~48 hr). For transformation, cells were diluted with YS media to OD_{600nm} ~0.16-0.18 and grown until OD_{600nm} reached ~0.20 (2-3 hr). 50 mL of cells were removed and placed on ice for 30 minutes. Cells were washed with 20 mM ice-cold sucrose as in Schleper et al., 1992. After the final wash, cells were resuspended in 400 mL ice-cold 20 mM sucrose and kept on ice. 100 μL of cells were added to a chilled 0.1 cm gap length cuvette (Bulldog Bio) and 2 μL of SSV DNA (100-500 ng/μL) was added to the cells. SSV DNA used in transformation was typically purified from *E. coli* using the GeneJET plasmid purification kit (Fisher). Cells were transformed by electroporation (BioRad Gene Pulser II) under the following conditions: 1.5 kV, 400 Ω, 25 μF. Immediately following electroporation, cells were resuspended in 1 mL of 70º C YS, transferred to a 1.5 mL tube, and incubated for 1 hr in a 70º C dry incubator. Following incubation, cells were transferred to 50 mL of preheated YS in a long neck Erlenmeyer flask and grown with shaking at 70º C. Cultures should appear turbid after ~24-36 hours.

Halo assay to check for infectious virus production

Following electroporation of *Sulfolobus* with SSV DNA, halo assays were performed in duplicate 48 and 72 hours post-transformation. Halo assays were performed as previously described (Chapter 2). Indicator lawns were always

prepared using the same strain of uninfected *Sulfolobus* that was used in the original transformation.

Confirmation of infectious SSV DNA

Transformed cultures that resulted in production of halos were further analyzed to confirm the identity of the viral DNA and ensure that contamination with a different virus had not occurred. Viral DNA was purified from infected cells using GeneJET kit (detailed above). Purified DNA was then PCR-amplified using DreamTaq DNA polymerase (Thermo-Fisher) and primers that flank the mutated region of the viral DNA. Control reactions using the original mutant DNA (i.e. DNA used in transformation of *Sulfolobus*) and wild type SSV1 DNA were always performed. Cultures harboring infectious SSV1 mutants were spun down, concentrated 10X in *Sulfolobus* storage buffer (YS medium with 20% glycerol, lacking yeast extract and sucrose, pH = 5.0), and frozen at -80º C.

Transmission electron microscopy

Samples were prepared for TEM analysis as in Iverson and Stedman 2012 (Chapter 2). Alternatively, samples were also prepared by directly spotting 5 μL of infected culture onto the TEM grid in place of supernatant.

Table 3-2: LIPCR primers used to construct SSV1 mutants

Mutation	Forward primer	Reverse Primer	Amino acids remaining in ORF[AB] / % of ORF remaining[AB]
ΔC102a	**CTG AAT GGC TAA AAA GAA CGG** (62)	**TGT TAT TAT TTC TGT TAC TGA GAC C** (57)	12 / 11%
ΔB251	**ATT GAC GAC GTA ACA AGA TAG** (56)	**TAA CCA TAA CCA TTC ATT ACT C** (55)	23 / 9%
ΔE54	**ACT CAT TTG TCC ACC TTG** (56)	**CAA CCA TAA** TAC TGT GAG G (53)	13 / 24%
ΔF92	**CTC TGA GTT GAA TAT CAT** TTT CC (58)	**ATA GCA TGG CTA GAA TAC AAG G** (59)	16 / 17%
ΔF93	**CAT** ATC CTC CTC ACT CCT CAG (60)	**ATC AAA AAG AGG TGA** ACT AGA TGG (61)	5 / 5%
ΔE96	**CAG ATC GAA TAT AGG AAC TTG C** (59)	**TAA** ATG ATA GAG AAG AGG AAA GAT AGG (60)	10 / 10%
ΔF112	CAT CTT GTA TGA ATT TAG AGT TTG TGC (63)	**AAG GCA AAG CAG TGA** ACT GAC (63)	14 / 13%
ΔB49	**GTA GAA GCA ATA AAT GAT TTG** (53)	**CTC AGA TTT TGC ACA TCC** (56)	15 / 30%
ΔF55	**TTT CCT CGG CAT** ACG CTA TC (64)	**TAA** ATG CCC TAC TAT ACT CTA TCT CTC TC (60)	4 / 7%
ΔA100	**TTT GAC TTC TGA** GGA GG (53)	**TAA CTC TTC TTC TTT TCG GG** (58)	15 / 15%
ΔC80	**TTA GCG AGG TAT GTA GAA AAT GTT TAG** ATG C (67)	**TGT GTA ACA TCT AGG TAA TTT GAT GTA TTC** (62)	23 / 29%
ΔA79	**GTT GAG TGA** ATA ATG TAT CAA TGT CTA C (60)	**GCA TCT AAA CAT** TTT CTA CAT ACC TC (60)	6 / 8%
ΔA45	**GTG AGA GGA CAA TGA** ACC (55)	**TTG ATA CAT** TAT TCA CTC AAC C (56)	7 / 15%
ΔC102b	**GGA TGA** CGG AGT CAG ACG TTG (67)	**GTT CAT** TGT CCT CTC ACC CTG AAG (67)	3 / 3%
B129 (ΔN-terminus)	**AAA GCG ATT TCA CAG TTT GTC** (61)	**TGA CTC CGT CAT** CCT CTA AC (59)	64 / 50%
B129 (ΔC-terminus)	**AGT TAG** GCT CTT TTT AAA GTC TAC C (58)	**TGA AAT CGC TTT ACT CGC** (59)	74 / 57%
ΔC124	**AGA AGA TAG** CCC TTT TTA AAG CC (62)	**GAA AAT AGA ACC TAC AAC TGT AAA CAG** (59)	14 / 11%
ΔB115	**TGG AGG GGT TTA AAA ACG TAA G** (63)	**TCA T**TC CGA CCC CCT AAT TAA C (65)	33 / 29%
ΔVP1	**TGA** GGG ATG GAA ATC AGT TTA AAG (64)	**CAA** ACT CCT TAG GAG TCT CAT CC (62)	2 / 1%
VP1(ΔN-terminus)	*ATG* **GAA GCA ACC AAC ATA GG** (61) and **GAA GCA ACC AAC ATA GG** (54)	**CAA** ACT CCT TAG GAG TCT CAT CC (62)	73 / 53% or 74 / 54%
VP1(E66A)[C]	**GCA GCA ACC AAC ATA GGC** (62)	**ACC TTT TGT GAG CTT GGG G** (63)	NA

VP1(E66Q)[D]	**CAA GCA ACC AAC ATA GGC** (59)	**ACC TTT TGT GAG CTT GGG G** (63)	NA
VP1 (Δ61-65)	**GAA GCA ACC AAC ATA GGC** (58)	**GGG GTT TGC CTT TGC TAC** (62)	133 / 96%
ΔVP3	**TGA** TAT GAA GTG GGT GCA AAA GG (67)	**CAT** CCC TCA CAC CTC AGT CTT TG (67)	2 / 2%
ΔD335 (Integrase)	**CAT** TTC GCC TCA CAG TAT TAT GG (64)	**GTC TGA** CAT TAC CCG TAT CAC (58)	2 / 1%
ΔD63-F92	**TTC TTT ACT CAT** TGT TTT TCA CCT TAG (62)	**ATA GCA TGG CTA GAA TAC AAG G** (59)	NA

❖ All sequences are written 5' → 3' orientation, bases in bold type denote those which are within the SSV1 ORF

A. Does not include the stop codon

B. Determined after deletion with LICPR

C. Highlighted base indicates mismatch that correlates to Glu→Ala mutation in VP1 protein

D. Highlighted base indicates mismatch that correlates to Glu→Gln mutation in VP1 protein

Results

To better understand the function of SSV1 genetic elements and extend the work presented in chapter 2, specific and random mutagenesis of SSV1 was performed:

Construction of specific SSV1 mutants by LIPCR

Deletions and point mutations were introduced into the SSV1 genome via LIPCR using Phusion High-Fidelity DNA polymerase (Thermo-Fisher). The majority of deletion mutants were constructed in the genetic background of pAJC97, an SSV1 shuttle vector with the TOPO pcr Blunt II plasmid (Thermo-Fisher) inserted into ORF *e178* (Clore and Stedman 2006; Table 3-3). Later, shuttle vectors EAI228 and EAI284 were also used as templates for LIPCR (See below and Table 3-3).

Template DNA for LIPCR was purified from 1.5 mL of a 5 mL overnight *E. coli* culture via alkaline lysis. Template DNA purified using a commercially available kit (Thermo-Fisher Plasmid Purification Kit) did not result in a noticeable improvement in PCR efficiency. LIPCR reactions (20 μL) were prepared following the DNA polymerase manufacturer's protocol. The only exception was that the recommended concentration of polymerase was reduced from 0.02 U/μL to 0.005 U/μL which helped to eliminate "smearing" that appeared using the higher concentration of enzyme. The optimal concentration of template DNA for LIPCR was determined empirically for each set of primers (See methods).

Primers were designed to delete a specified ORF while preserving the sequences of overlapping ORFs, promoter elements, and other genetic elements (Reiter et al., 1987b; Palm et al., 1991; Fröls et al., 2007; Fusco et al., 2013). Optimally, only the start and stop codons of the targeted ORF would remain following LIPCR, however, this was not always achievable. The majority of deletions resulted in ORFs ~10 codons in length (Table 3-2). The annealing temperature for each set of primers was estimated using primer design software (see methods). This estimate was often not optimal but was typically within 1-5˚ C of the optimal temperature.

In initial PCR experiments, an extension of 30 sec/kb at 72º C was used and resulted in little to no amplification or the appearance of large smears on the gel. The extension time was incrementally decreased to 18.75 sec/kb, which, along with the reduced enzyme concentration, eliminated smearing and more

consistently resulted in full-length product. A total of 35 amplification cycles were performed, although later experiments indicated that 30 cycles were sufficient.

Following LIPCR, phosphorylated DNA was ligated and transformed into *E. coli* for selection of full-length mutant DNA. Several commercial *E. coli* strains were tested for their ability to efficiently uptake ligation reactions, including chemically competent DH5α (NEB), DH5α (Invitrogen), 10-Beta (NEB), and NovaBlue Singles (Millipore). The only strain that could be reliably transformed were NovaBlue *E. coli* that were prepared in house (see materials and methods; Table 3-1). Transformants harboring full length pAJC97-derived deletion mutants were typically slow growing, formed small colonies on LB-Kan plates, and did not grow well in liquid media. This was likely due to the large size of the plasmid (~19 kb) and the high copy number inherent to TOPO pcr Blunt II. Later it was observed that pir^{+} *E. coli* (Epicentre) grew much more efficiently when transformed with EZTn5-based SSV1 shuttle vectors, most likely due to the reduced copy number of the plasmid. A few of the later SSV1 mutants were constructed in the genetic background of EAI283 and EAI228, which harbor the EZTn5 transposon in ORFs *e178* and *f112*, respectively (Tables 2 and 3).

Table 3-3: List of plasmids and SSV1 shuttle vectors used in this work

Plasmid	Description	Infectious in S441 (+ or -)	Reference
pAJC96	pAJC97 background with Integrase (*d335*) deleted	–	Clore and Stedman 2006
pAJC97	SSV1 shuttle vector (TOPO PCR Blunt II inserted into ORF *e178* at bp 3,173)	+	Clore and Stedman 2006
REC228 [A]	SSV1::Tn5 mutant, EZ-Tn5 inserted at bp 4,680 (ORF *f112*)	+	This Work
REC229 [A]	SSV1::Tn5 mutant, EZ-Tn5 inserted at bp 11,265 (*vp4*)	–	This Work
REC230 [A]	SSV1::Tn5 mutant, EZ-Tn5 inserted at bp 8,277 (ORF *a291*)	–	This Work
REC231 [A]	SSV1::Tn5 mutant, EZ-Tn5 inserted at bp 11,227 (*vp4*)	–	This Work
EAI232	SSV1::Tn5 mutant, EZ-Tn5 inserted at bp 8,662 (ORF *a291*)	–	This Work
EAI239	SSV1::Tn5 mutant, EZ-Tn5 inserted at bp 5,218 (ORF *b49*)	–	This Work
EAI240	SSV1::Tn5 mutant, EZ-Tn5 inserted at bp 4,792 (ORF *f112*)	–	This Work
EAI241	SSV1::Tn5 mutant, EZ-Tn5 inserted at bp 100 (ORF *a100*)	–	This Work
EAI242	SSV1::Tn5 mutant, EZ-Tn5 inserted at bp 10,159 (*vp4*)	–	This Work
REC243 [A]	SSV1::Tn5 mutant, EZ-Tn5 inserted at bp 14,046 (ORF *c84*/*a82*)	–	This Work
REC244 [A]	SSV1::Tn5 mutant, EZ-Tn5 inserted at bp 4,788 (ORF *f112*)	–	This Work
REC245 [A]	SSV1::Tn5 mutant, EZ-Tn5 inserted at bp 12,718 (*vp1*)	–	This Work
EAI247	SSV1::Tn5 mutant, EZ-Tn5 inserted at bp 13,992 (ORF *c84*/*a82*)	–	This Work
EAI248	SSV1::Tn5 mutant, EZ-Tn5 inserted at bp 953	–	This Work
EAI249	SSV1::Tn5 mutant, EZ-Tn5 inserted at bp 6,375 (ORF *a132*)	+	This Work
EAI250	SSV1::Tn5 mutant, EZ-Tn5 inserted at bp 11,641 (ORF *b78*)	–	This Work
EAI251	SSV1::Tn5 mutant, EZ-Tn5 inserted at bp 14,677 (ORF *b277*)	–	This Work
EAI253	SSV1::Tn5 mutant, EZ-Tn5 inserted at bp 3,889 (ORF *e51*)	+	This Work
EAI254	SSV1::Tn5 mutant, EZ-Tn5 inserted at bp 8,633 bp (ORF *a291*)	–	This Work
EAI255	SSV1::Tn5 mutant, EZ-Tn5 inserted at bp 7,509 (ORF *b129*)	–	This Work
EAI256	SSV1::Tn5 mutant, EZ-Tn5 inserted at bp 2,776 (ORF *d*244)	+	This Work
EAI257	SSV1::Tn5 mutant, EZ-Tn5 inserted at bp 4,249 (ORF *e96*)	+	This Work

EAI258	SSV1::Tn5 mutant, EZ-Tn5 inserted at bp 8,998 (ORF *c124*)	+	This Work
EAI260	SSV1::Tn5 mutant, EZ-Tn5 inserted at bp 80 (ORF *a153*)	–	This Work
EAI261	SSV1::Tn5 mutant, EZ-Tn5 inserted at bp 14,209 (ORF *b277*)	–	This Work
REC262	SSV1::Tn5 mutant, EZ-Tn5 inserted at bp 4,209 (ORF *e96*)	+	This Work
EAI266	SSV1::Tn5 mutant, EZ-Tn5 inserted at bp 1,988 (ORF *e54*)	–	This Work
EAI267	SSV1::Tn5 mutant, EZ-Tn5 inserted at bp 1,717 (*Integrase*)	–	This Work
EAI271	SSV1::Tn5 mutant, EZ-Tn5 inserted at bp 6,018 (ORF *a100*)	–	This Work
EAI278	SSV1::Tn5 mutant, EZ-Tn5 inserted at bp 14,834 (ORF *b277*)	–	This Work
EAI281	SSV1::Tn5 mutant, EZ-Tn5 inserted at bp 13,709 (ORF *c84*/*a82*)	–	This Work
EAI282	SSV1::Tn5 mutant, EZ-Tn5 inserted at bp 11,807 (ORF *c166*)	–	This Work
EAI283	SSV1::Tn5 mutant, EZ-Tn5 inserted at bp 3,572 (ORF *e178*)	+	This Work
EAI286	SSV1::Tn5 mutant, EZ-Tn5 inserted at bp 11,407 (*vp4*)	–	This Work
EAI296	SSV1::Tn5 mutant, EZ-Tn5 inserted at bp 5,783	–	This Work
EAI297	SSV1::Tn5 mutant, EZ-Tn5 inserted at bp 12,170 (ORF *c166*)	–	This Work
EAI305	SSV1::Tn5 mutant, EZ-Tn5 inserted at bp 7,359 (ORF *c102b*)	–	This Work
EAI319	SSV1::Tn5 mutant, EZ-Tn5 inserted at bp 7,387 (ORF *c102b*)	–	This Work
REC322 [A]	SSV1::Tn5 mutant, EZ-Tn5 inserted at bp 5,573 (ORF *f55*)	–	This Work
REC324 [A]	SSV1::Tn5 mutant, EZ-Tn5 inserted at bp 4,394 (ORF *d63*)	+	This Work
REC325 [A]	SSV1::Tn5 mutant, EZ-Tn5 inserted at bp 967 (*Integrase*)	+	This Work
EAI446	SSV1::Tn5 mutant, EZ-Tn5 inserted at bp 13,211 (*vp3*)	+[B]	This Work
EAI452	SSV1::Tn5 mutant, EZ-Tn5 inserted at bp 13,003	+[B]	This Work
EAI453	SSV1::Tn5 mutant, EZ-Tn5 inserted at bp 5,451	+	This Work
EAI469	SSV1::Tn5 mutant, EZ-Tn5 inserted at bp 13,491 (*vp2*)	–	This Work
EAI476	SSV1::Tn5 mutant, EZ-Tn5 inserted at bp 13,191 (*vp3*)	+[B]	This Work
EAI477	SSV1::Tn5 mutant, EZ-Tn5 inserted at bp 5,247 (*b49*)	–	This Work

EAI486	SSV1::Tn5 mutant, EZ-Tn5 inserted at bp 5,024	+	This Work
EAI492	SSV1::Tn5 mutant, EZ-Tn5 inserted at bp 3,837 (*f92*)	+	This Work
EAI202	pAJC97 background with ORF *b129* deleted	–	This Work
EAI201	pAJC97 background with ORF *d244* deleted	–	This Work
EAI205	pAJC97 background with ORF *b49* deleted	+	This Work
EAI206	pAJC97 background with ORF *b251* deleted	–	This Work
EAI214	pAJC97 background with ORF *b115* deleted	–	This Work
EAI216	pAJC97 background with ORF *e96* deleted	+	This Work
EAI233	pAJC97 background with ORF *a100* deleted	–	This Work
EAI237	pAJC97 background with N-terminus of *vp1* deleted	–	This Work
EAI327	pAJC97 background with ORF *f112* deleted	–	This Work
EAI390	pAJC97 background with ORF *c124* deleted	+	This Work
EAI394	pAJC97 background with ORF *a79* deleted	+	This Work
EAI398	pAJC97 background with ORF *a45* deleted	+	This Work
EAI400	pAJC97 background with ORF *f55* deleted	+	This Work
EAI407	pAJC97 background with ORF *f92* deleted	+	This Work
EAI413	pAJC97 background with ORF *f93* deleted	+	This Work
EAI420	pAJC97 background with *vp3* deleted	+	This Work
EAI421	pAJC97 background with ORF *c102a* deleted	–	This Work
EAI422	pAJC97 background with C-terminus of ORF *b129* deleted	–	This Work
EAI427	pAJC97 background with E66Q mutation in VP1	+	This Work
EAI430	pAJC97 background with N-terminus of *b129* deleted	–	This Work
EAI435	pAJC97 background with ORF *e54* deleted	–	This Work
EAI439	pAJC97 background with ORF *c80* deleted	+	This Work
EAI496	pAJC97 background with ORF *c102b* deleted	–	This Work
EAI499	pAJC97 background with ORF *b129* deleted	–	This Work

EAI500	pAJC97 background with E66A mutation in VP1	+	This Work
EAI553	EAI283 background with SSV9-*vp1* complementation	+	This Work
EAI557	EAI283 background with N-terminus of SSV9-*vp1* complement	+	This Work
EAI561	EAI283 background with SSV9-*b279* complementation	–	This Work
EAI563	EAI283 background with SSV9-*a231* complementation	–	This Work
EAI564	EAI283 background with N-terminus of *vp1* deleted	–	This Work
EAI566	EAI283 background with SSV2-*vp1* complementation	+	This Work
JAH572	SSV1::Tn5 mutant, EZ-Tn5 inserted at bp 5,016	–	This Work
JAH573	SSV1::Tn5 mutant, EZ-Tn5 inserted at bp 5,264 (*b49*)	–	This Work
JAH576	SSV1::Tn5 mutant, EZ-Tn5 inserted at bp 5,681	–	This Work
EAI578	EAI283 background with residues 61-65 deleted from VP1	–	This Work
EAI580	EAI283 background with ORF *d335* deleted	–	This work
EAI582	EAI228 background with ORFs *d63* through *f92* deleted	+	This work

[A.] Plasmids with REC prefix refer to SSV1 mutants that were isolated in the Recombinant DNA Techniques Laboratory at Portland State University

[B.] Insertions in these mutants were removed in *Sulfolobus* via homologous recombination

Table 3-4: List of other primers used in this work

Primer	Sequence	Description
Kan-2 FP-1	ACCTACAACAAAGCTCTCATCAACC	Used to sequence transposon mutants, anneals ~100 bp upstream EZ-Tn5 5' end
Univ_7F	ATTCAGATTCTGWATWCAGAA	Amplifies structural gene region of all SSVs with universal 8R (Snyder et al., 2004)
Univ_8R	TCSCCTAACGCACTCATC	Amplifies structural gene region of all SSVs with universal 7F (Snyder et al., 2004)
SSV9_VP1_F	GAAGTTTGGTCAAAGTTAAACG	Amplifies VP1 gene from SSV9 for complementation of SSV1-ΔVP1
SSV9_VP1_R	ATCTTTGTAGATTTTATACG	Amplifies VP1 gene from SSV9 for complementation of SSV1-ΔVP1
SSV2_VP1_F	GCCACCAGACTAATGCTAAGC	Amplifies VP1 gene from SSV2 for complementation of SSV1-ΔVP1
SSV2_VP1_R	GTCACGATATATCTTATACGCTATGAC	Amplifies VP1 gene from SSV2 for complementation of SSV1-ΔVP1
SSV9_VP1NT_F	GAAGTTTGGTCAAAGTTAAACG	Amplifies N-terminus of VP1 gene from SSV9 for complementation of SSV1-VP1ΔN-terminus
SSV9_VP1NT_R	ACCCCTAGTAAGTTTGGG	Amplifies N-terminus of VP1 gene from SSV9 for complementation of SSV1-VP1ΔN-terminus
B277_LIPCR_F	AATGGTCTCAGTAACAGAAATAATAAC	Deletion of entire *b277* ORF from EAI283 for complementation
B277_LIPCR_R	GCAAACACCTCAACCCAAG	Deletion of entire *b277* ORF from EAI283 for complementation
SSV9_C279_F	ATGTCTGATGGAAAATTAGTTTCC	Amplifies c279 ORF from SSV9 for complementation of Δ*b277* mutant
SSV9_C279_R	TCACCCCTTACTCGTATTTGC	Amplifies c279 ORF from SSV9 for complementation of Δ*b277* mutant
B251_LIPCR_F2	TGATACGGGTAATGTCAGACC	Deletion of entire *b251* ORF from EAI283 for complementation
B251_LIPCR_R2	GGTCAAATAATTGTTTCAGAATTG	Deletion of entire *b251* ORF from EAI283 for complementation
SSV9_A231_F	ATGGAAATGACTGAACTATTATGG	Amplifies A231 ORF from SSV9 for complementation of Δ*b251* mutant
SSV9_A231_R	CTACTCTAAATTTTCGTTATTTTTCC	Amplifies A231 ORF from SSV9 for complementation of Δ*b251* mutant

* All primer sequences written 5' → 3' orientation

Construction of SSV1 EZTn5 transposon insertion mutants

The EZ-Tn5^{TM} <R6Kγori/KAN-2> transposon insertion kit (Epicentre) was used to construct insertion mutants of SSV1. Combining EZTn5 transposon DNA and SSV1 DNA at an equimolar ratio as specified by the manufacturer yielded ~5% of transformants harboring full length shuttle vectors. By increasing this ratio to 30:1 (SSV1 DNA:EZTn5 DNA) the number of transformants containing full length plasmid increased to ~50%. SSV1::EZTn5 constructs could only be isolated from transformation of "low copy number" pir^{+} *E. coli (*Epicentre; Table 3-1) but could not be isolated following transformation of "high copy number" pir-116 *E. coli* (Epicentre; Table 3-1).

Summary of SSV1 mutagenesis

All 35 ORFs in the SSV1 genome were mutated via insertion and/or deletion and these mutants were tested for the ability to infect *S. solfataricus* S441 cells, a host permissive to SSV1 infection (Stedman unpublished). Additional insertion mutants within specific intergenic regions were also isolated and characterized. The infectivity of SSV1 mutants was assayed by transformation of mutant DNA into uninfected S441 cells and then spotting these cultures on an uninfected lawn 72 and 96 hours later. Spotting a productively infected culture on a lawn of cells results in the appearance of a ring of growth inhibition, i.e. a "halo", due to the production of virus by the infected cells. A large proportion of our SSV1 mutants did not result in the production of halos

when transformed, indicating they are not infectious. However, a false negative may result from a failed transformation and this is indistinguishable from the phenotype of a non-infectious mutant. To help rule out this possibility, a mutant was only labeled as non-infectious after a minimum of five independent negative results had been documented with appropriate negative and positive controls. While this does not completely eliminate the possibility of false negatives, it does significantly increase the likelihood that a non-infectious phenotype is due to the mutation and not an experimental error. Electroporation data from EAI262, which was used as a positive control in most experiments, show that 97% of transformations result in production of virus as indicated by halo assay. However, we encountered two mutants that were initially judged to be non-infectious after five negative results only to later find that they were capable of producing infectious virus (EAI258 and EAI453).

Overall SSV1 appears quite tolerant of mutagenesis, allowing for a variety of insertions and deletions throughout the genome (Table 3-3; Figure 3-2). With a few exceptions, ORFs that could be deleted from the virus without loss of infectivity also tolerated insertion of the 2 kb EZTn5 transposon. The extent to which a particular ORF was conserved within the *Fuselloviridae* generally dictated whether or not the ORF was essential for SSV1 infectivity. Of the 12 ORFs that are unique to SSV1, 8 of these could be mutated without a loss of infectivity. Conversely, there are 12 universally conserved genes belonging to

the so-called fusellovirus core (Table 1-2), none of which could be mutated with the notable exception of the *vp3* gene (Figures 3-1 and 3-2).

The moderately conserved ORFs of the T6 transcript

The ~2 kb region of the SSV1 genome encoding the T6 transcript (Reiter et al., 1987b) encompasses seven ORFs (*a100, a132, c80, a79, a45, c102b,* and *b129*), most of which are moderately-to-highly conserved within the *Fuselloviridae* (Figure 3-1). ORF *a132* is unique to SSV1 and apparently non-essential as it was the only ORF within this region that tolerated insertion of the transposon. All three of the ORFs *c80*, *a79*, and *a45* could be deleted without a loss of infectivity. Oddly, infectious virus was not produced when the transposon was inserted into ORFs *a45* and *c80*, potentially due to polar effects caused by the insertion. The highly conserved ORFs *c102b* and *a100* were found to be essential for infectivity. ORF *b129* was previously determined to be essential for SSV1 infectivity (Chapter 2).

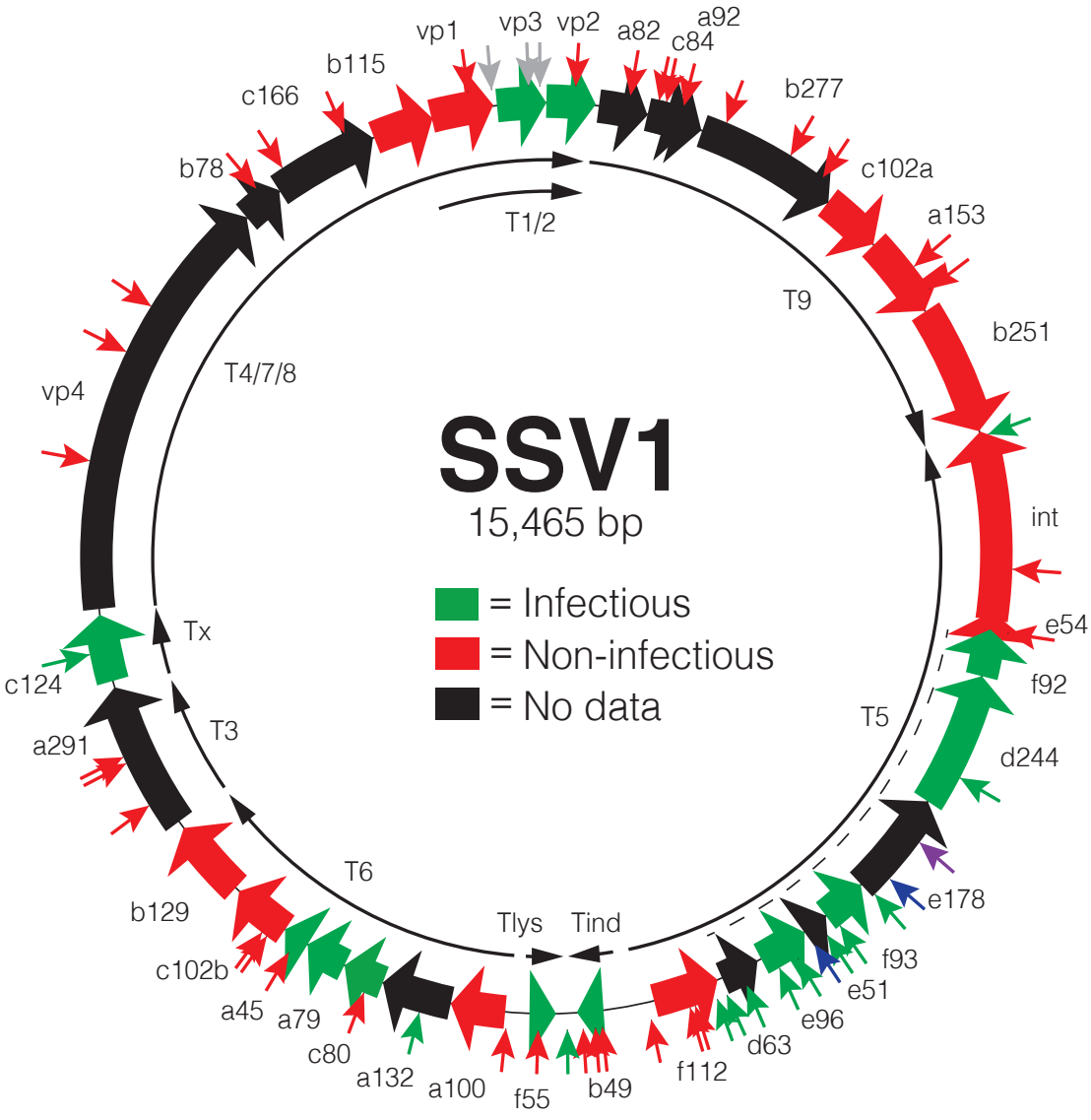


Figure 3-2: Results of SSV1 mutagenesis
Open reading frames are displayed and labeled as in Figure 3-1. The fill color of the ORF indicates if deletion of the ORF results in production of infectious virus (green) or not (red). ORFs for which there are no deletion data are black. The dotted line near the T5 transcript indicates the region deleted in the infectious mutant EAI582. Arrows on the exterior of the map denote the location of individual transposon insertions and the color of the arrow indicates whether the insertion results in the production of infectious virus (green) or not (red). Grey arrows in the *vp1*/*vp3* gene region indicate insertions that are apparently removed via homologous recombination, allowing for the production of infectious virus. Blue arrows (ORFs e51 and e178) denote infectious SSV1 insertion mutants reported by Stedman et al., 1999. The purple arrow in ORF e178 denotes the location of pBluescript insertion in the infectious mutant pKMSD48 (Stedman et al., 1999) and also denotes the location of the Topo pcr blunt II plasmid in infectious mutants pAJC97 and pAJC96 (Clore and Stedman 2006). Transcripts are denoted by small curved arrows on the interior of the map and labeled as in Figure 3-1.

Both the N- and C-termini of B129 are essential for SSV1 infectivity

The structure of SSV1-B129 displays N-terminal and C-terminal DNA binding domains, both of which have been shown separately to bind DNA (Lawrence et al., 2009 and unpublished). Mutants lacking the *b129* N- and C-termini were constructed via LIPCR and tested for infectivity (Figure 3-3A). Neither mutant appears capable of producing infectious virus, implying that the full length B129 protein is required for infectivity.

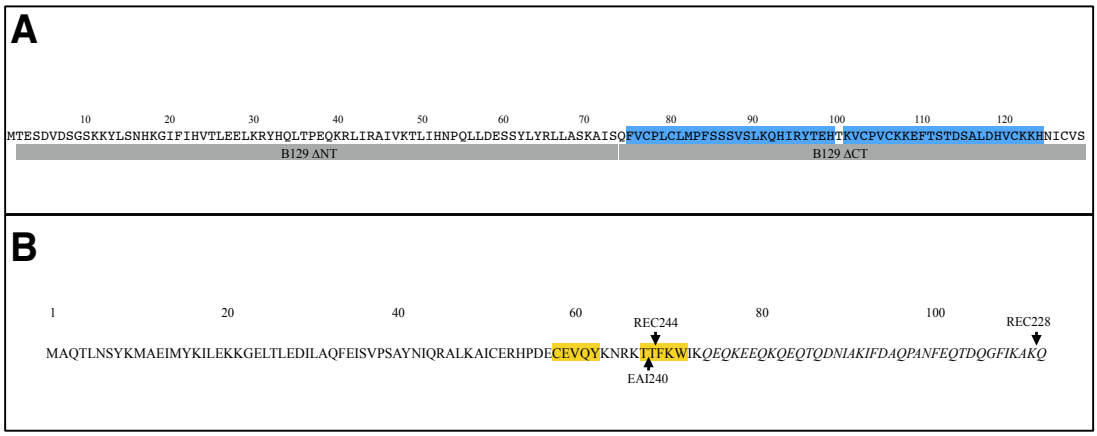


Figure 3-3: B129 and F112 protein sequences
(A) Amino acid sequence of the SSV1-B129 protein (NP_039795.1). Residues that were deleted to construct the *b129* N- and C-terminal deletions are labeled with grey bars under the sequence. Blue highlighted residues indicate the position of the tandem ZNF motifs in the C-terminus (Lawrence et al., 2009).
(B) Amino acid sequence of the SSV1-F112 protein (NP_039787.1). Location of the three transposon insertion mutations have been mapped to the F112 sequence and are indicated by arrows. Residues that correspond to the "wing" motif in the F112 crystal structure are highlighted in yellow and C-terminal residues not present in the structure are italicized (Menon et al., 2008).

The monocistronic transcripts T3 and Tx

The monocistronic transcripts T3 and Tx are both expressed early in the transcription cycle (Fröls et al., 2007). The T3 transcript encodes ORF *a291* and Tx encodes ORF *c124*, both of which reportedly have homologues in SSV2

(ORFs *305* and *126*, respectively) but were not identified here using pBLAST software (NCBI) (Fusco et al., 2015c; Figure 3-1).

Three separate insertion mutants (EAI230, EAI232, and EAI254) in *a291* were analyzed, none of which resulted in the production of infectious virus (Figure 3-2). Conversely, *c124* tolerated insertion in and deletion of the entire ORF without the loss of infectivity. It should be noted that the *c124* insertion mutant (EAI258) produced infectious virus in only three of twelve independent transformations, something that was not observed for other mutants nor for the *c124* deletion mutant (EAI390).

The fusellovirus core is intolerant of mutagenesis

The fusellovirus core genes are predominantly encoded on one half of the SSV1 genome and encompass the middle-to-late transcripts T4/7/8, T1/2, and T9 (Fröls et al., 2007; Figure 3-1). Core genes *c166*, *b115*, *a82*, *c84*, *a92*, *b277*, *a153*, and *b251* did not tolerate deletions and/or insertions (Figures 3-1 and 3-2). This was expected due to their presumed significance in the fusellovirus life cycle (Redder et al., 2009; Chapter 2). ORFs *vp4* and *b78*, both proposed to encode the SSV1 tail filament (Redder et al., 2009; Quemin et al., 2015), likewise appeared essential for infectivity. The only two ORFs in this region that are not well conserved are *c102a* and the structural gene *vp2* which was previously shown to be non-essential (Iverson and Stedman 2012; Figure 3-1). ORF *c102a* could not be deleted and is apparently essential for infectivity. Unexpectedly,

SSV1 lost infectivity when the transposon was inserted into the *vp2* gene, in disagreement with the previous deletion data (Chapter 2).

The T5 transcript region is almost entirely dispensable

The T5 transcript region encodes a number of poorly conserved ORFs (Figure 3-1) which probably accounts for why this region is extremely tolerant of mutagenesis. Ten ORFs occupy this region, seven of which were shown to be non-essential (Figure 3-2). ORF *f112* appeared to be essential as it could not be deleted without loss of infectivity. However, ORF *f112* tolerated transposon insertion in the extreme C-terminus of the ORF (REC228) but did not tolerate either of two insertions further upstream (EAI240 and EAI244; Figure 3-3B). Mapping of the insertion mutants to the structure of the F112 protein shows that in the functional mutant REC228, the transposon is inserted at the end of the C-terminal unstructured region of the protein whereas the two non-infectious mutants (EAI240 and EAI 244) have the transposon inserted within the beta-sheets that contribute to the "wing" of F112 (Menon et al., 2008).

ORF *d244* was previously deleted and shown to be non-essential (Chapter 2). In agreement with this, two *d244* insertion mutants yielded functional virus. *S. solfataricus* GΘ cells infected with Δ*d244* virus exhibit significant growth retardation compared to cells infected with wild-type virus (Chapter 2). Interestingly, host growth retardation was not observed in S441 or GΘ cells following infection with either of the *d244* insertion mutants (Data not

shown). Comparison of D244 homologues in the *Fuselloviridae* has indicated that D244 is probably mis-annotated as it contains a 32 residue N-terminal extension that is absent from all other homologues (Lawrence, personal communication). An alternative ATG start codon is located 96 bases downstream from the original start codon and produces a better alignment with other fusellovirus homologues over the entirety of the protein (data not shown).

The large stretch of non-essential ORFs occupying the T5 transcript hint that the majority of this region may not be required for SSV1 infectivity. LIPCR was used to delete the 2.4 kb region encompassing ORFS *f92*, *d244*, *e178*, *f93*, *e51*, *e96*, and *d63* in the genetic background of REC228 (Figure 3-2, dotted line). Somewhat surprisingly, SSV1 remained infectious after the loss of almost 15% of its genome, raising the question of why SSV1 devotes a significant portion of genetic information to apparently dispensable genes.

Integrase *and* e54 *deletion mutants exhibit a variable host range*

The *integrase* gene and ORF *e54* lie adjacent to each other at the distal end of the T5 transcript region (Figure 3-1). SSV1 lacking the integrase gene (pAJC96, Table 3-3) was previously shown to be capable of infecting *S. solfataricus* P2 (Clore and Stedman 2006). In this study, the integrase deletion mutant was unable to infect strain S441 but was capable of infecting strain GΘ (Table 3-2). Similarly, SSV1 with a deletion of ORF *e54* was unable to infect S441 but able to infect GΘ. A new *integrase* deletion was constructed (EAI580)

in the EAI283 genetic background that has the same host-range phenotype as the original mutant pAJC96. A mutant with the transposon inserted in *e54* was unable to infect S441 nor GΘ. An insertion in the integrase gene likewise inhibited SSV1 infectivity, although a mutant (EAI325) with an insertion near the C-terminus was capable of producing virus in S441 (Figure 3-2).

ORFS b49 *and* f55 *and the predicted origin of replication*

The 1.3 kb region between the T5 and T6 transcripts harbors two poorly conserved ORFS (*b49* and *f55*), the putative origin of replication, and several promoters (Cannio et al., 1998; Figures 3-1 and 3-4). ORF *b49* is encoded on the UV-inducible transcript T_{ind} while *f55* is encoded on the recently discovered transcript T_{lys} (Reiter et al., 1987b; Fröls et al., 2007; Fusco et al., 2013). It was observed that both of these ORFs could be deleted from SSV1 without a loss of infectivity, although insertion was not tolerated by either ORF (Figures 3-2 and 3-4). Three *b49* insertion mutants were analyzed (EAI239, EAI477, and JAH573), all of which failed to yield infectious virus. Likewise, the *f55* insertion mutant (EAI322) was unable to produce infectious virus. Transposon insertions in the intergenic region surrounding these ORFs were also analyzed (Figure 3-4; Table 3-3). Insertion of the transposon between ORFs *b49* and *f55* (EAI453) did not inhibit virus infectivity whereas insertion downstream of ORF *f112* (EAI572) and upstream of ORF *a100* (EAI296) eliminated infectivity. In EAI572, the transposon is inserted two base pairs from the transcription factor B recognition

element of the T5 promoter and may explain why the mutant is non-functional. In EAI296, the insertion occurs at the base pair following the mapped transcription start site for the T6 promoter. Furthermore, both insertions in EAI572 and EAI296 fall within an 11 bp sequence that is repeated several times in this region, weakly binds the *f55* putative transcription regulator, and has been implicated in SSV1 replication (Fusco et al., 2013; Qureshi 2007; Figure 3-4).

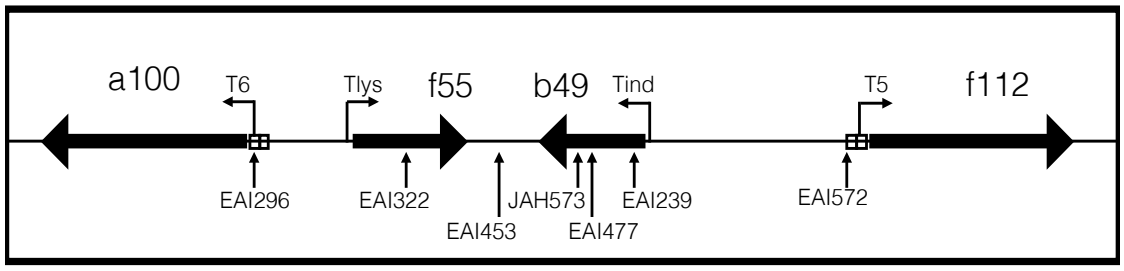


Figure 3-4: The T5, T6, T_{ind}, and T_{lys} promoter region
The region spanning bases 4600-6100 of the SSV1 genome is displayed. ORFs are shown as block arrows and labeled. The locations of transposon insertions are indicated by arrows and labeled (See Table 3-3). The mapped T5, T6, T_{lys} and T_{ind} promoters are labeled and the transcription start sites are indicated by bent arrows (Reiter et al., 1987b; Fröls et al., 2007; Fusco et al., 2013). Sequence repeat elements overlapping the T5 and T6 promoters and involved in STRIP binding are indicated by boxes (Qureshi 2007).

The minor capsid gene vp3 is highly conserved but non-essential

The structural genes *vp1* and *vp3* are highly conserved and hypothesized to be essential for SSV1 infectivity (Figure 3-1). Surprisingly, SSV1 remained infectious following deletion of the *vp3* gene. The Δ*vp3* mutant was also shown to be infectious in *Sulfolobus solfataricus* strains GΘ, P1 and P2 suggesting this is not a strain-specific phenomenon (Table 3). Because it is not uncommon for *Sulfolobus* to encode cryptic fusellovirus genomes/genes, it is possible that the Δ*vp3* mutant was able to remain infectious via complementation. The sequence of GΘ is unavailable but the P1 and P2 genomes have been sequenced and do

not contain any obvious *vp3* homologues. A more likely explanation is that VP1 may be compensating for the absence of VP3 due to the homology that exists between VP3 and the C-terminus of VP1 (Figure 3-7B). TEM images of Δ*vp3* virions shows that although the particles are spindle-shaped, they are much more elongated relative to wild-type (Figure 3-5). Nonetheless, it appears that *vp3* is non-essential for viral infection despite its high degree of conservation and being a structural protein in the fusellovirus virion (Reiter et al., 1987a; Menon et al., 2008; Quemin et al., 2015).

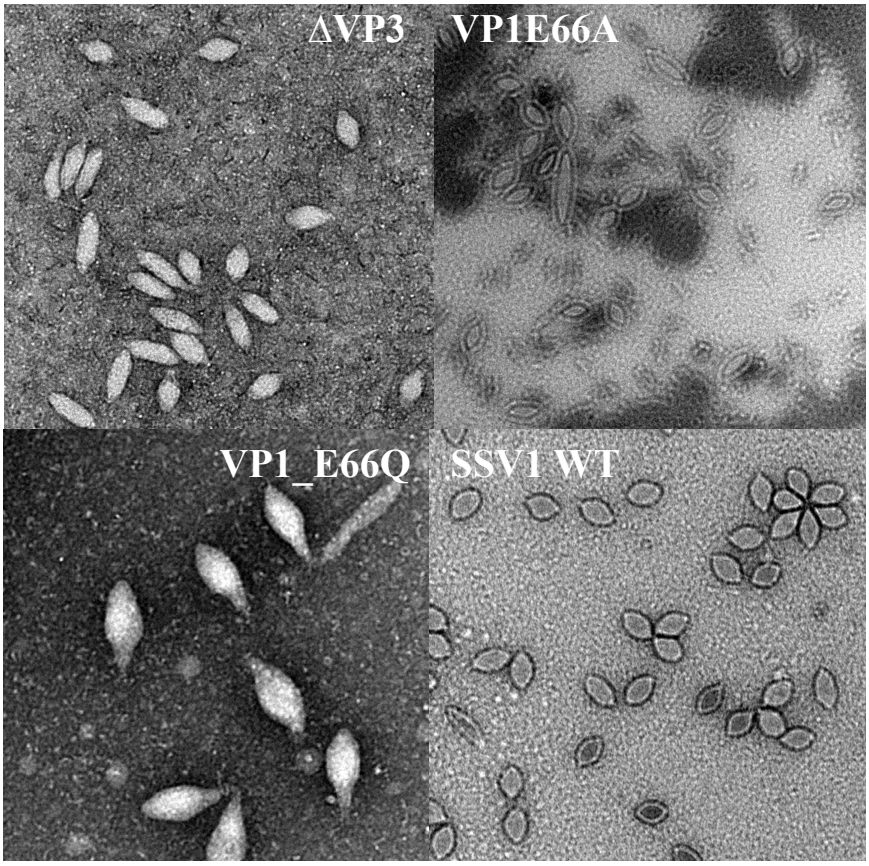


Figure 3-5: Electron micrographs of SSV1 structural mutants
Transmission electron micrographs of purified SSV1 wild-type, -Δ*vp3* (EAI420), -VP1E66A (EAI500), and -VP1E66Q (EAI427) virions are displayed. See methods for purification protocol. SSV1 WT particles are ~76 x 40 nm, ΔVP3 particles are ~120 x 26 nm, ΔVP1E66Q particles are ~ 93 x 42 nm, and VP1E66A particles are ~ 91 x 34 nm (measurements determined by averaging at least 25 particles).

Homologous recombination can result in deletion of the vp3 *gene*

Two mutant constructs with transposon insertions in the *vp3* gene (EAI446 and EAI476) and one mutant with an insertion in the small intergenic region between *vp1* and *vp3* (EAI452) were shown to generate infectious virus (Figure 3-6A). Following infection, viral DNA was purified and screened via PCR using primers that amplified the region encompassing the *vp1*, *vp3*, and *vp2* structural genes (Table 3-4). Oddly, each mutant yielded truncated PCR products compared to the wild-type SSV1, indicating that a deletion had occurred within this region instead of the transposon insertion (Figure 3-6B). There is an identical 61 bp sequence in the C-termini of both the *vp1* and *vp3* genes (Figure 3-6C; Palm et al., 1991). DNA sequencing of each PCR product show that almost the entire *vp3* gene was missing. Apparently a deletion occurred, facilitated by homologous recombination between the 61 bp repeats (Figure 3-6D). Recombination also resulted in a deletion of the final 15 bp from the *vp1* gene, including the stop codon. The new *vp1* stop codon is supplied by the native *vp3* stop codon (Figure 3-6D). Presumably similar recombination events occur in wild type SSV1 DNA, however this has not been detected to date. Recombination is thought to have occurred in *Sulfolobus* following transformation. PCR of the full-length mutant DNA from the transposition reaction did not produce the truncated product that is observed in the DNA purified from *Sulfolobus* following electroporation (Figure 3-6B). Moreover, the

pir^{+} *E. coli* used here have the *rec*A1 genotype which should inhibit their ability to undergo recombination (Table 3-1).

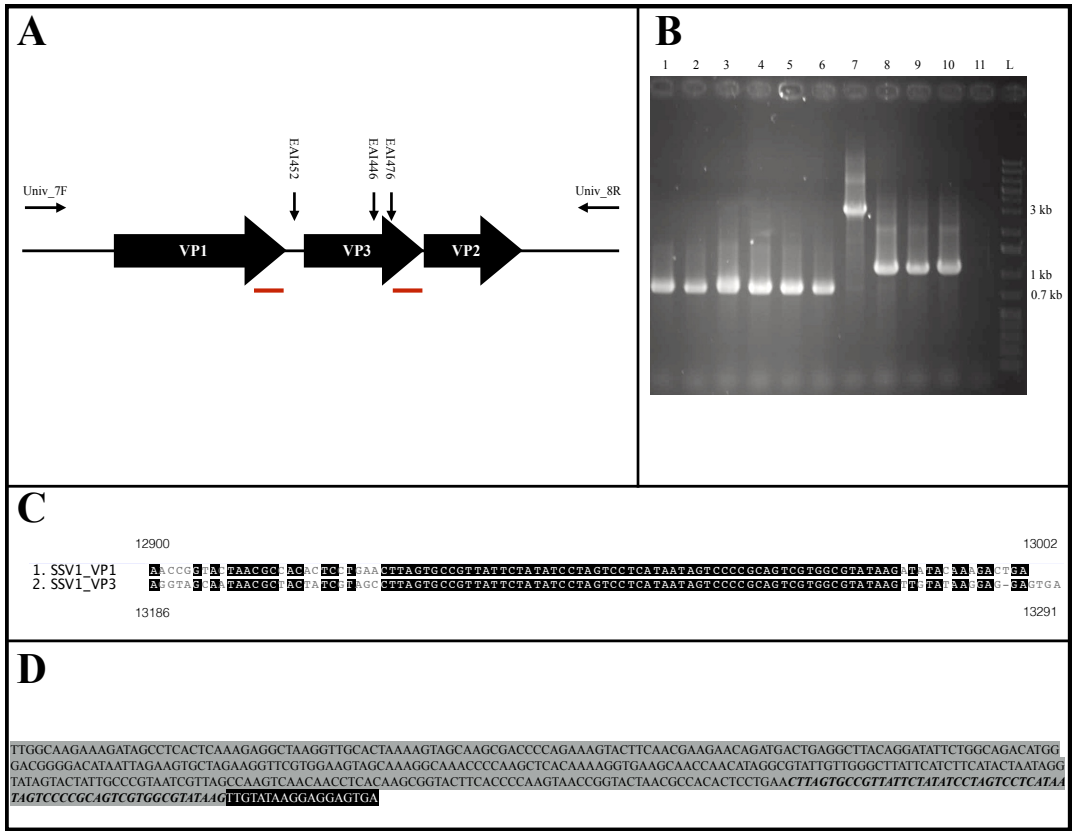


Figure 3-6: Analysis of functional insertion mutants EAI446, EAI452, and EAI476 within VP1/VP3 structural gene region

(A) The SSV1 structural genes *vp1*, *vp3* and *vp2* are displayed. PCR primer (univ_7F and univ_8R) annealing sites are indicated by thin horizontal arrows. Vertical arrows denote the location of the transposon is in each mutant. Red underlined regions indicate 61 bp repeats in the 3' end of the *vp1* and *vp3* genes.

(B) PCR of viral DNA purified from infected *Sulfolobus* S441 with primers univ_7F and univ_8R to confirm transposon insertion within the structural gene region. Lanes 1 and 2 are from EAI446; Lanes 3 and 4 are from EAI452; Lanes 5 and 6 are from EAI476. Lane 7 is the original EAI446 DNA (from *E. coli*) used to transform S441. Lanes 8-10 are wild-type SSV1 DNA. Lane 11 is a no template negative control. Lane 11 is GeneRuler 1 kb Plus Ladder (Fisher), relevant molecular weights are indicated beside the gel.

(C) Pairwise alignment of C-termini from *vp1* and *vp3* genes displaying the 61 bp identical sequence encoded by both genes. Identical bases are highlighted in black. Position of displayed region in the SSV1 genome is indicated in the figure.

(D) Nucleotide sequence of the structural gene region following recombination in mutants EAI446, EAI452, and EAI476. Grey highlighted bases correspond to *vp1* gene; black highlighted bases correspond to remainder of *vp3* gene. Italicized bases denote the 61 bp repeat that presumably facilitates recombination.

Mutagenesis of the major capsid gene vp1

As expected, the *vp1* gene was found to be essential due to its prominent role as the major capsid protein. A complete deletion of *vp1* and an insertion in the middle of the ORF both failed to yield infectious virus (Figure 3-2). The VP1 protein appears to be proteolytically cleaved at an internal completely-conserved glutamate residue to produce the mature protein found within the virion (Reiter et al., 1987a; Quemin et al., 2015; Figure 3-7). To investigate if the encoded N-terminus is required for infectivity, this region was deleted while leaving the conserved glutamate intact (Figure 3-7). Infectious virus was not produced by the *vp1*-ΔN-terminus mutant (EAI564), suggesting that the VP1 N-terminus is required for production of infectious virus. The N-termini of fusellovirus VP1 proteins are less conserved than the C-termini, with the exception of a patch of well-conserved residues just upstream of the conserved glutamate (Figure 3-7). This region was deleted from the *vp1* gene and the resulting mutant (EAI578) was similarly found to be non-infectious. This reinforces the idea that the N-terminus of VP1 is essential and the conserved residues near the cleavage site may be important for proteolysis.

We also investigated whether the universally conserved glutamate of VP1 is essential for infectivity or if it could be replaced. A single base pair change was made in the *vp1* gene by LIPCR, converting the glutamate to the structurally similar but functionally distinct glutamine (Table 3-2). Interestingly, this mutation did not impair SSV1 infectivity (Figure 3-7). A second mutation in which the

glutamate was converted to the less structurally conserved alanine likewise failed to impair virus infectivity. Virions of both *vp1* point mutants were analyzed by TEM and shown to yield abnormal particles (Figure 3-5). Together, these data demonstrate that the glutamate residue is not essential for SSV1 infectivity despite its universal conservation in the *Fuselloviridae*. It is unknown if proteolysis of VP1 can still occur at this position in either of these mutants or at one of the other well conserved glutamate residues upstream (Figure 3-7).

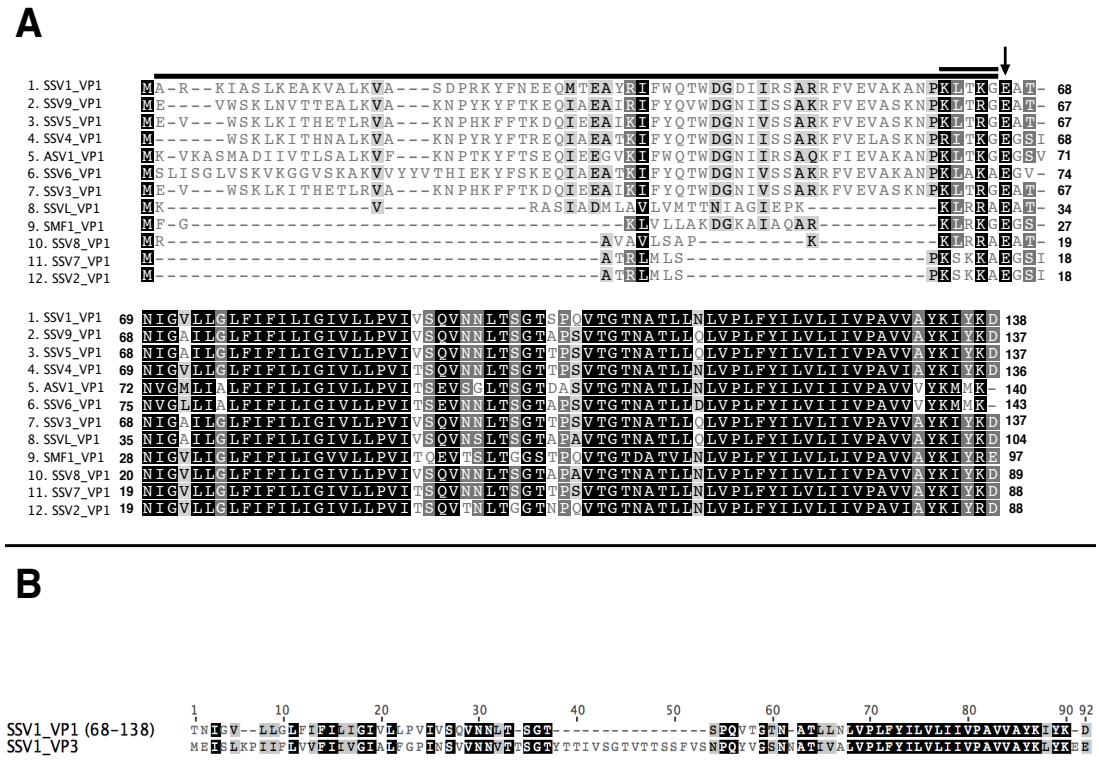


Figure 3-7: Fusellovirus VP1 and VP3 proteins
(A) Multiple sequence alignment of predicted VP1 proteins from 12 fusellovirus genomes (Table 1-1). Residues highlighted in black are identical, residues highlighted in grey are conserved, and residues that are not highlighted are not conserved. Residues deleted to construct *vp1*-ΔN-terminus mutant (EAI237, EAI564) are indicated by longer line at the top of the alignment. Well conserved residues 61-65 deleted from mutant EAI578 are indicated by the shorter line above the alignment. The universally conserved glutamate that was mutated to a glutamine (EAI427) and alanine (EAI500) is indicated by an arrow. Alignment was made with default parameters using Geneious (Biomatters) software.
(B) Pairwise alignment of SSV1-VP1 and SSV1-VP3 proteins displays SSV1-VP3 is homologous with the C-terminus of SSV1-VP1. Conserved bases are highlighted in black, semi-conserved bases are highlighted in gray and non-conserved bases are not highlighted.

Complementation in cis of SSV1 deletion mutants

In order to confirm that SSV1 deletions were non-functional only due to the deletion, we complemented the mutants in cis with homologs from other SSVs. The Δ*vp1* mutants could be rescued when complemented in cis with the homologous *vp1* genes from SSV9 or SSV2 (Figure 3-7). The C-termini of the three VP1 proteins are highly homologous, however, the N-terminus of SSV2-VP1 is considerably shorter (15 residues) than the other two (SSV9-VP1 65 residues, SSV1-VP1 66 residues). The non-functional *vp1*-ΔN-terminus mutant (EAI564) was also rescuable by complementation in cis with the similar SSV9 *vp1*-N-terminus (Figure 3-7). Similar attempts were made to complement two other non-infectious SSV1 deletion mutants (Δ*b277* and Δ*b251*). Both genes were replaced by their putative homologs, SSV1-Δ*b277* with SSV9-*b279* and SSV1-Δ*b251* with SSV9-*a231*. However, functional virus was not produced.

Discussion

Overview of SSV1 mutagenesis

The genome of SSV1, like all archaeal viruses, encodes a large number of ORFs that share little to no similarity with sequences in public databases, confounding investigations of fusellovirus life cycles. In light of this, we have conducted a genetic analysis of each of the ORFs in the fusellovirus SSV1. Our findings illustrate that the SSV1 genome is surprisingly malleable and encodes a significantly higher number of non-essential genes than previously thought (Stedman et al., 1999). Using a combination of transposon mutagenesis and LIPCR, we have shown that 16 of the 35 ORFs can be disrupted without a loss of infectivity. This is somewhat surprising because viral genomes are often (with some exceptions) streamlined to encode only the minimal genes needed for survival (Cann 2016). As expected, almost all SSV1 ORFs that are well conserved among the *Fuselloviridae* appear to be essential for infectivity while less conserved ORFs were more tolerant of mutation.

A significant amount of research on SSV1 genetics has been performed over the last decade or so. Genetic analysis of SSV1, the first performed on any archaeal virus, provided the first clues into which ORFs were essential for SSV1 (Stedman et al., 1999). Stedman et al., inserted the bacterial plasmid pBluescript into various SSV1 ORFs following partial restriction endonuclease digestion and the resulting mutants were assayed for infectivity essentially as here. Insertions of pBluescript into ORFs *e178* and *e51* were tolerated, whereas insertions into

ORFs *e96*, *b129*, *vp4*, and *d335* (Integrase) all failed to produce infectious virus (Figure 3-2). These data are in agreement with the findings of this work, with the exception of *e96*, which was shown to be non-essential by both deletion and insertional mutagenesis (Figure 3-2). Polar effects caused by the insertion of pBluescript may explain this discrepancy, although this seems unlikely as we have not only isolated a number of infectious insertion mutants in this region, we have also shown that this entire quadrant of the SSV1 genome can be disposed of without loss of infectivity. Alternatively, there could have been problems with the original transformation into *Sulfolobus*, something that has been encountered in our research (e.g. infectious mutants EAI258 and EAI453 were initially judged to be non-infectious).

The T5 transcript is almost entirely dispensable

The T5 transcript is expressed early in the SSV1 transcription cycle (along with T6) and encodes some of the least conserved fusellovirus ORFs (Fröls et al., 2007; Figure 3-1). Only three of the ten T5 ORFs appear to be essential for infectivity, and two of these (*integrase* and *e54*, see below) appear to be essential only in specific hosts (Figure 3-2). Despite it being non-conserved, *f112* was the only ORF in this region essential for infectivity. The structure of F112 displays a winged helix-turn-helix DNA binding motif, however a DNA binding sequence has not been identified (Menon et al., 2008). Similar to B129, F112 possesses an intramolecular disulfide bond that was experimentally shown

to improve protein stability (Menon et al., 2008). The role of F112 in the fusellovirus life cycle is unknown, but the gene, or an unidentified genetic element in this region, is apparently critical for SSV1 infectivity (Figure 3-2).

After it was observed that so many of the ORFs in this region appear non-essential, a deletion mutant lacking ORFs *f92*, *d244*, *e178*, *f93*, *e51*, *e96*, and *d63* was constructed (~2.4 kb) and found to be infectious. This result suggests that SSV1 devotes at least 15% of its genome to genes/ORFs that are seemingly superfluous, although it is unknown what effect this deletion has on the long-term fitness of SSV1 or in the environment. This seems like an odd strategy although it is not completely unheard of in the viral world, as evidenced by analyses of mycobacteriophage genomes (Hatfull 2015).

Work on mycobacteriophages has revealed some interesting parallels with the *Fuselloviridae* and may help explain these results (Hatfull 2015). Mycobacteriophage genomes contain a set of well-conserved structural/assembly genes but much of their genomes are composed of small ORFs (~500 bp) of unknown function and whose presence seems to vary considerably from isolate to isolate (Hendrix et al., 1999). Knockouts of these genes were constructed and roughly 2/3 of the non-structural/assembly genes were not essential to mycobacteriophage function, very reminiscent to what is seen in SSV1 (Marinelli et al., 2008). One hypothesis is that these genes may have been required for growth in an ancestral host but are surplus to requirements in the current host. Alternatively, the natural phage environment is

very competitive and it has been observed that phage encode restriction modification systems as well as components of the CRISPR-Cas defense system in an effort to out-compete invading phage and/or for protection against host defense systems (Hatfull 2015). Therefore, it is possible that some of these small genes of unknown function may play a role in this process. It is unknown if any fusellovirus genes are involved in a "defensive" role, although many fuselloviruses (not SSV1) do encode a Cas4 homologue (Appendix B).

The moderately conserved ORFs of the T6 transcript

The T6 transcript, which is expressed early in the SSV1 transcription cycle, encodes a number of well conserved ORFs whose products have been shown via structural studies or bioinformatics to exhibit DNA binding motifs that are characteristic of transcription regulators (Fröls et al., 2007; Prangishvili et al., 2006; Lawrence et al., 2009). The ORFs encoding three of these putative transcription factors (*c80*, *a79*, and *a45*) could be deleted without a loss of infectivity. However, neither *c80* nor *a45* tolerated insertion of the transposon. Insertion of the transposon was by and large not tolerated anywhere in this region, with the exception of ORF *a132* which is apparently unique to SSV1. ORFs *a100* and *c102b*, whose functions remain mysterious, were also found to be essential in accordance with their high level of conservation in the *Fuselloviridae* (Figure 3-1).

Previous work showed that the fusellovirus core ORF *b129*, which appears to have N- and C-terminal DNA binding domains, was essential for infectivity (Chapter 2; C.M. Lawrence personal communication). Here we showed that deletion of either DNA binding domain resulted in a loss of infectivity, suggesting that both are required for SSV1 function. The structure of the B129 dimer (PDB ID: 2WBT) indicates that the N-terminal alpha-helical region of B129 forms the dimer-dimer interface, and these interactions appear to be destroyed by deletion of the N-terminus. In addition to the two tandem ZNF domains, the C-terminus of B129 also encodes an intramolecular disulfide bond that is hypothesized to improve protein stability (Menon et al., 2008).

The SSV1 integrase *gene may be essential in some hosts*

The *integrase* gene was the first deletion made in SSV1 and showed that the integrase gene was not essential for infectivity, although it was quickly out-competed by wild type virus (Clore and Stedman 2006). However, this mutant (pAJC96 a.k.a. Δ*int*) could not infect the *Sulfolobus solfataricus* strain used herein (S441), which is slightly different from the Δ*int*-susceptible host strain (P2) used in the original work. The Δ*int* virus was able to infect *Sulfolobus solfataricus* strain GΘ, indicating that the Δ*int* mutant has a variable host range. *Sulfolobus* genomes are known to encode a number of integrase-like genes, one of which may be capable of complementing the Δ*int* virus (She et al., 2001a; She et al., 2004). The genome of GΘ has not been sequenced but likely encodes

integrase-like genes. If complementation occurs, this might imply that a similar phenomenon occurred in P2 during the original SSV1-Δ*int* work. Although it was shown that Δ*int* virus had not integrated in its natural location in the P2 genome by PCR, integration facilitated by an exogenous integrase could have occurred at a different site (Clore and Stedman 2006). *Sulfolobus* is also host to a plethora of extrachromosomal elements, at least one of which (pSSVi) is known to encode an SSV-like tyrosine recombinase (Wang et al., 2007; Wang et al., 2015). We have not identified any extrachromosomal element in the strains used here, however, it is not unimaginable that such an element may be present and has the potential to complement the *integrase* deletion in the *Sulfolobus* strains GΘ and P2 (Clore and Stedman 2006). In any event, the findings presented here question the conclusion that SSV1 does not require an *integrase* gene for infectivity.

Deletion of ORF *e54*, which is non-conserved and lies just upstream of the integrase start codon, resulted in the same phenotype as the Δ*int* mutant. It is possible that deletion of *e54* somehow disrupts expression of the integrase gene, effectively resulting in a double mutant. The *integrase* gene occupies the 3' end of the T5 transcript and is believed to be transcribed mainly via the T5 promoter. Although it has not been identified, there is evidence that the integrase gene may actually be transcribed from its own promoter. *Integrase* mRNA was found in greater abundance than T5 mRNA following UV induction and a similar phenomenon was observed for the *integrase* gene during an analysis of the

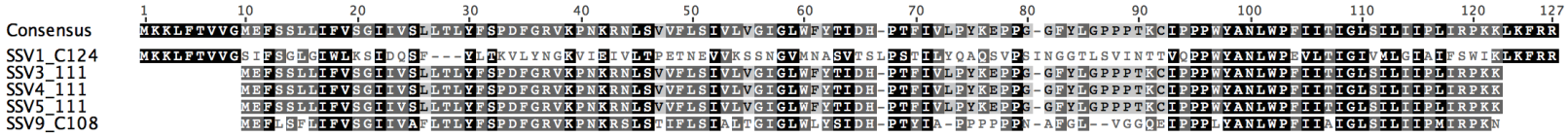
SSV2 transcription cycle (Fröls et al., 2007; Ren et al., 2013). Although an obvious promoter has not been identified upstream of the SSV1 *integrase* gene, several fuselloviruses (SSV2, SSV3, SSV4, and SSV9) do encode putative promoters upstream of the integrase gene (Clore 2008). It is worth noting that ORF *f92*, which also occupies this region of the SSV1 genome, can be deleted without loss of infectivity (Figure 3-2). Based on the low conservation of the *e54* ORF and the evidence for an *integrase* promoter in this region, it seems likely that deletion of *e54* simultaneously disrupts expression of the *integrase* gene and explains the identical phenotype exhibited by both mutants.

The Tx and T3 transcripts

The monocistronic transcripts T3 and Tx encode ORFs *a291* and *c124*, respectively. *c124* was shown to be non-essential for SSV1 infectivity while *a291* is apparently essential (Figure 3-2). The T3 and Tx transcripts, and the homologous transcripts in SSV2 (encoding ORFs *305* and *126*, respectively), were found to be constitutively expressed in cells infected with either virus (Fusco et al., 2015c). Interestingly, most of the other fusellovirus genomes appear to encode similarly sized ORFs in an identical position just upstream of the putative tail fiber gene (*vp4*). Despite little to no detectable homology among these ORFs via a pBLAST homology search, there is evidence that transcription of these ORFs is conserved in the *Fuselloviridae* (Ren et al., 2013).

A pairwise alignment of SSV1-C124 and SSV2-126 showed limited identity (12%) exists between the two. Alignment of all putative C124-like proteins (i.e. those occupying an analogous position in the genome) indicates SSV1-C124 is homologous to four other fusellovirus proteins (SSV3-111, SSV4-111, SSV5-111 and SSV9-C108) especially in the C-terminus (Figure 3-8A). SSV8 possesses two short ORFs (*b74* and *c82*) in the analogous genomic region whose products are homologous to one another and also to the products of ORFs in SSV7 (ORF *67*) and SSVL (ORF *75*) (Figure 3-8B). A homologue for SSV2-126 could not be identified. Each of these ORFs is preceded by a putative promoter that is similar to the Tx promoter of SSV1, suggesting that transcription of each of these ORFs may be conserved (data not shown).

A


Consensus
SSV1_C124
SSV3_111
SSV4_111
SSV5_111
SSV9_C108

B

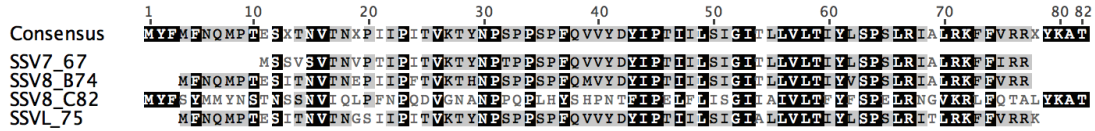

Consensus
SSV7_67
SSV8_B74
SSV8_C82
SSVL_75

Figure 3-8: Alignment of putative SSV1-C124 homologues occupying a similar position in the fusellovirus genome
Multiple sequence alignment of putative SSV1-C124 homologues (panel A) and alignment of proteins that occupy a similar position in the virus genome (i.e. downstream of *vp4*) but are not homologous to SSV1-C124 (panel B). SSV2-126 was not homologous to any fusellovirus protein (not shown). Conserved amino acids are highlighted in black, semi-conserved amino acids are highlighted in gray and non-conserved amino acids are not highlighted. A consensus sequence for each alignment was generated and is displayed at the top of the alignment. SSV8 encodes two homologous ORFs at this position in the genome, both of which are shown (panel B). SSV3-111, SSV4-111, and SSV5-111 are apparently identical in each virus. Alignments were made with default parameters using Geneious (Biomatters) software.

Pairwise alignment of SSV1-A291 and SSV2-305 revealed that the N-terminal 20 amino acids of the two proteins are highly similar, while the remaining protein exhibits little to no similarity (Figure 3-9A). Because similarity was not identified in the original pBLAST search, the remaining 10 fusellovirus genomes (Table 1-1, SMF1 was omitted) were re-examined and nine additional homologues of SSV1-A291 were identified (Figure 3-9A). Each of the A291-like proteins possesses a well conserved N-terminus but the remainder of the protein is non-homologous. Furthermore, each of these ORFs is preceded by a putative promoter that shares significant similarity with the T3 promoter in SSV1, indicating that transcription of these genes is probably conserved (Figure 3-9B). Because A291 appears to be well conserved in the *Fuselloviridae*, it is thus not surprising that it appears essential for SSV1 infectivity. These results show that SSV1-A291 and SSV1-C124 both have more homologues than previously detected, indicating that the degree of conservation for SSV1 proteins presented here should be considered a lower limit (Figure 3-1).

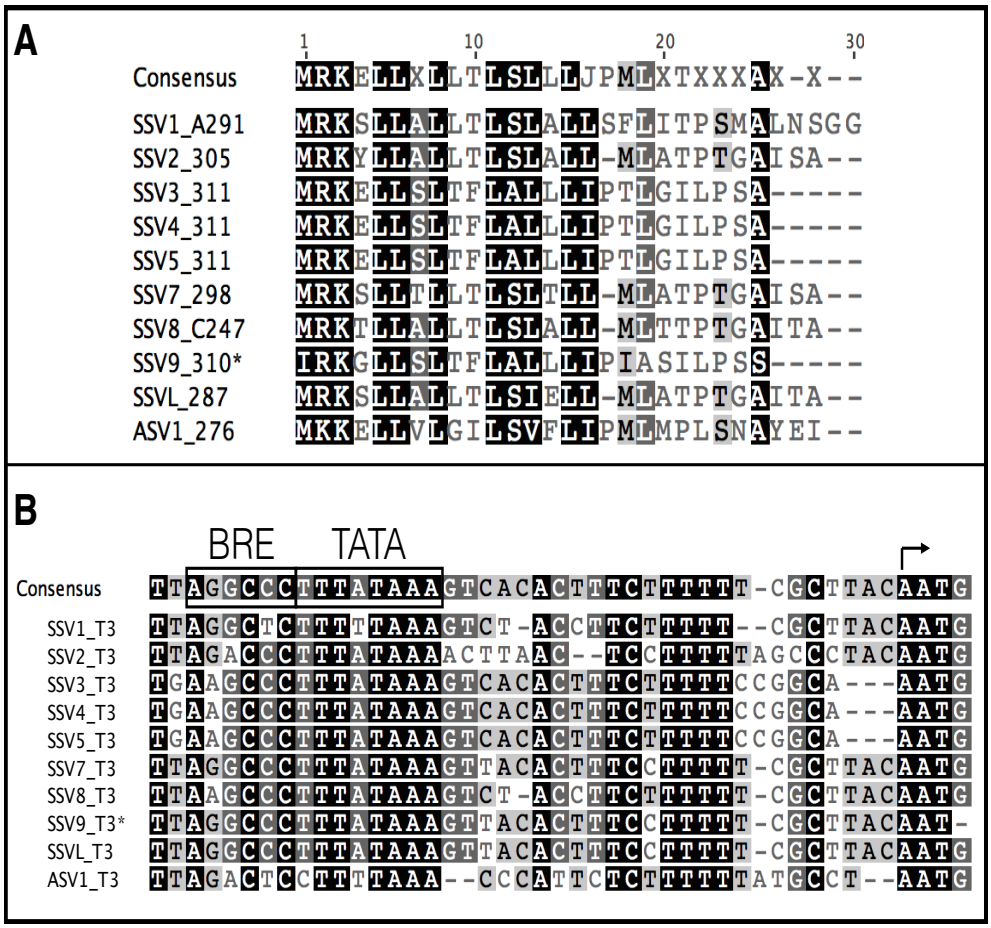


Figure 3-9: Conservation of SSV1-A291 and the T3 transcript in the *Fuselloviridae*
(A) Multiple sequence alignment of the N-termini from fusellovirus SSV1-A291 homologues. Conserved amino acids are highlighted in black, semi-conserved amino acids are highlighted in gray and non-conserved amino acids are not highlighted. A consensus sequence was generated and displayed at the top of the alignment. SSV9-310* is hypothesized to be the full length version of SSV9-B252 (NP_963990.1), however the hypothetical start codon contains a G→A missense mutation which results in a truncated product (B252). This mutation was confirmed by sequencing (data not shown). The putative promoter for SSV9 ORF *310* is apparently conserved (see panel B). Alignment was made with default parameters using Geneious (Biomatters) software.
(B) Multiple sequence alignment of the putative T3 promoter from SSV1 and putative homologues in other fuselloviruses. Universally conserved bases are highlighted black, highly-conserved bases are highlighted dark gray, semi-conserved bases are highlighted light gray and non-conserved bases are not highlighted. The consensus sequence for the promoter is displayed above the alignment. Predicted BRE and TATA elements (Fröls et al., 2007) are boxed and labeled, the mapped T3 transcription start site in SSV1 is indicated by a bent arrow (Reiter et al., 1987b; Fröls et al., 2007). See panel A for explanation of SSV9 ORF *310**. Alignment was made with default parameters using Geneious (Biomatters) software.

The role of A291 is unknown but there are clues about its function. The N-terminus may facilitate a well-conserved interaction with a host or viral protein. An intriguing hypothesis is that A291 may play a role in viral assembly due to the simultaneous upregulation of the T3 transcript with the structural genes *vp1*, *vp3*, and *vp4* (Fröls et al., 2007; Fusco et al., 2015c). However, A291 was not identified by mass spectrometry in purified SSV1 virions, suggesting that A291 is either not present in the mature virion or may be present at undetectable levels (Menon et al., 2008; Quemin et al., 2015). Perhaps A291 acts as a protein scaffold during assembly, a role which has not yet been assigned to any fusellovirus gene. The Tx transcript is upregulated concurrently with the T3 transcript, suggesting that its protein product (C124) could also be involved in assembly and may potentially interact with A291.

The fusellovirus core is intolerant of mutagenesis

The fusellovirus core is the set of genes/ORFs that are encoded by all known fuselloviruses. The core currently consists of 12 genes/ORFs (Figure 3-1; also see Table 1-2), almost all of which appear to be essential for SSV1 infectivity (Figure 3-2). This is not surprising and re-enforces the idea that these are critical to the viral lifecycle. The only non-essential core gene identified was *vp3* which is discussed below. Core fusellovirus ORFs are predominantly clustered in one half of the genome and are upregulated mid-to-late in the SSV1 transcription cycle (Fröls et al., 2007). Their timing of transcription and co-

expression with known structural genes hints that most of the core genes may have roles in virus replication, assembly, and packaging although there is no experimental evidence to support this. It is somewhat peculiar that we were not able to isolate a single functional insertion mutant within this entire half of the genome, even in the poorly conserved *vp2* gene which is known to be non-essential from deletion data (Chapter 2). This is likely an example of disruption caused by a polar effect and serves to remind that other apparently non-infectious mutants may only appear so due to secondary effects of the transposon and not due to the mutation of the ORF itself.

ORFs B49 and F55 and the predicted origin of replication

Unlike the other known fuselloviruses, transcription of SSV1 is strongly induced by UV irradiation and is highly temporally regulated (Fröls et al., 2007; Reiter et al., 1987b). Following UV irradiation, the transcript T_{ind} is immediately upregulated and is swiftly followed by upregulation of the two flanking transcripts T5 and T6 (Figure 2; Reiter et al., 1987b; Fröls et al., 2007). *b49* is the only ORF encoded by T_{ind} and possesses no homology to sequences in public databases, including the *Fuselloviridae* (Figure 3-1). Due to the abundance of T_{ind} immediately following UV irradiation, it seems likely that the B49 protein plays a role in activation of viral transcription, either directly or indirectly. We have shown that the *b49* ORF can be deleted and is apparently not essential for SSV1 infectivity. This agrees well with transcriptomic data from non-UV induced SSV1-

infected cells, where T_{ind} (and *b49*) was not detected and presumably not required for infection (Fusco et al., 2013). However, the effect of the *b49* deletion on the SSV1 response to UV irradiation is unknown and should provide insight into this unique mechanism.

SSV1 encodes a second monocistronic transcript in this region, T_{lys}, that is apparently expressed constitutively (Fusco et al., 2013; Fusco et al., 2015c). T_{lys} encodes a 55 amino acid protein (F55) that was shown experimentally to weakly bind (K_d ~3-10 μM) early viral promoters (T5, T6, T_{ind}, and T_{lys}) and is hypothesized to repress transcription from them during the maintenance of the carrier-state or "lysogeny" (Fusco et al., 2013). Our results show that the *f55* gene is not required for the production of infectious virus, but it is not clear what effect its absence has on SSV1. Based on the above hypothesis, absence of the F55 protein would result in a loss of repression of early viral promoters which could lead to constant expression of early gene products throughout the infection. Such an effect should be observable in the transcriptome of this mutant and in synchronous infections (see Chapter 4).

Microarray analyses of SSV1 and SSV2 lysogens (i.e. stably infected cells) of *S. solfataricus* LnF1 cells were recently conducted and it was observed that SSV1 has a very minimal impact on host gene expression whereas SSV2 results in significant changes to the host transcriptome, including activation of the CRISPR-Cas antiviral defense system (Fusco et al., 2015c). Interestingly, SSV1 and SSV2 express almost the same complement of genes throughout lysogeny,

one exception being the constitutive expression of *f55* by SSV1. This has led to speculation that F55 may play a key role in mitigating the host response to SSV1 infection, and experiments with our Δ*f55* mutant will be useful for investigating this hypothesis.

This area of the genome, which contains an abundance of promoters as well as the putative origin of replication, was very intolerant of transposon insertion (Figure 3-4). The only functional insertion mutant in this region (EAI453) contained the transposon between the *f55* and *b49* ORFs, a significant distance from any of the known regulatory elements (Figure 3-4). All other insertions in this region fall within one of the two ORFs or were located adjacent to a promoter. Because the deletion data showed that *b49* and *f55* are not essential, it is unclear why insertions within these ORFs do not produce functional virus. Unpublished data from the Bell laboratory has mapped the origin of replication to this region and a self-replicating vector containing the putative SSV1 origin was reportedly constructed, although this has been difficult to reproduce in other laboratories (Fröls et al., 2007; Cannio et al., 1998; Stedman, personal communication). GC and purine skew analyses also indicated that the origin is within this area and appears well conserved in other fuselloviruses (Clore 2008). A protein complex known as STRIP (SSV-1 T5/T6 Region Interacting Protein) was isolated from *S. shibatae* (the original SSV1 host) and shown to bind tandem repeat sequences located upstream of the T5 and T6 promoters, although the significance of this was not determined (Qureshi

2007; Figure 3-4). Interestingly F55 has also been shown to weakly bind these same sequences, suggesting a related function between the two.

The vp3 *minor capsid gene is not essential*

The only non-essential core gene was *vp3*, which came as a major surprise considering its high degree of conservation and presence as the minor structural protein within the virion. VP3 and the major capsid protein VP1 were both isolated from purified SSV1 virions, although VP3 was reportedly found at a lower abundance (Reiter et al., 1987a; Quemin et al., 2015). Because the VP3 and proteolytically processed VP1 proteins are highly homologous (Figure 3-7B), we hypothesize that VP1 complements the Δ*vp3* mutant. The genomes from two of the three strains susceptible to infection with the Δ*vp3* mutant (P2 and S441) were scanned for *vp3* homologues using BLAST software (NCBI). No homologues were detected indicating that complementation from the host chromosome is unlikely. Δ*vp3* virions were examined by TEM and are elongated relative to wild-type SSV1 (Figure 5). Elongated particles are often observed in the virions of SSV9, SSV6, and ASV1, however, each of these viruses encodes a VP3 homologue (Wiedenheft et al., 2004; Redder et al., 2009). Nonetheless, the dispensability of a seemingly critical gene is unexpected and the consequences of its loss in regards to virion stability and structure remain to be investigated.

Following the finding that the Δ*vp3* virus was infectious, several mutants harboring transposon insertions within the *vp1*-*vp3* gene region were isolated

and assayed for infectivity (Figure 3-6). Unlike other insertion mutants in the core region of the genome, insertions within *vp3* and the *vp1*/*vp3* intergenic space seemed to be tolerated and did not interfere with the production of infectious virus. However, when the viral DNA was purified from infected cultures it was observed that a deletion instead of an insertion had occurred and this appears to have been facilitated by homologous recombination between two identical 61 base pair sequences at the C-termini of the *vp1* and *vp3* genes. Recombination also results in the near complete deletion of the *vp3* gene (Figure 3-6C and 3-6D). Full length viral DNA harboring the transposon could not be recovered from infected *Sulfolobus* cells, indicating that the transposon disrupts SSV1 infectivity and must be eliminated to produce infectious virus albeit at a loss of the *vp3* gene in the process.

Our data suggest that this recombination event occurs within *Sulfolobus* following transformation of the full length mutant DNA and not in *E. coli* prior to transformation. If this is the case, it seems likely that the same recombination occurs in wild-type SSV1 populations, resulting in sporadic loss of the non-essential *vp3* gene. We have not been able to identify or isolate a spontaneous *vp3* mutant in WT SSV1 populations, suggesting that these mutants may be present at very low abundances and/or they are outcompeted by wild-type virus. SSV1 is the only fusellovirus that possesses such a large stretch of 100% identical bases within the *vp1* and *vp3* genes. The next largest stretch of 38 bases occurs in SMF1 *vp1*/*vp3* genes while the rest of the fuselloviruses do not

encode stretches longer than 25 base pairs and most do not encode direct repeats longer than 12 base pairs (data not shown).

Mutagenesis of the major capsid gene vp1

Unsurprisingly, the major capsid protein (MCP) VP1 is essential for SSV1 infectivity. All fusellovirus VP1 proteins are homologous in their C-termini and encode a universally conserved glutamate residue hypothesized to be involved in proteolysis to yield the mature VP1 protein found within the virion (Reiter et al., 1987a; Quemin et al., 2015; Figure 3-7). The source of the protease is unknown (viral or host) and the process might also be autocatalytic. Despite the universal conservation of the glutamate residue, we have shown that substitution with a glutamine or alanine residue is tolerated although the effect of these substitutions on VP1 maturation is unknown (Figure 3-7). We have also shown that despite the absence of the N-terminus in the mature VP1 protein, this region appears to be essential and could not be deleted without a loss of infectivity. Furthermore, deletion of the five amino acids immediately upstream of the conserved glutamate also result in a loss of infectivity. These residues are very well conserved in fusellovirus VP1 proteins and appear to be critical for infectivity (Figure 3-7).

Proteolytic processing during virion assembly is quite common and may actually be the rule rather than the exception (Dougherty and Semler 1993). Recent work with the MCP gp5 from the bacteriophage HK97 reveals strong

similarities with VP1 from SSV1 (Oh et al., 2014). Assembly of HK97 particles begins with the association of hexamers and pentamers of the MCP and a phage encoded protease (gp4) to form the immature procapsid (Duda et al., 1995a). Similar to VP1, gp5 possesses an N-terminal extension (the Δ-domain) that is cleaved during assembly and is absent in the mature virion (Duda et al., 1995b; Oh et al., 2014). Deletion of the N-terminal Δ-domain (completely or partially) from gp5 abolishes HK97 infectivity, similar to what is observed in SSV1. The Δ-domain also seems to act as a chaperone, as gp5 lacking this region was shown to be very unstable and formed insoluble aggregates that were not observed with wild-type protein. The promotion of assembly and stability along with its absence from the mature virion implies that the Δ-domain has adopted the role of a scaffold, a protein that is absent from HK97 (Duda et al., 1995b). Furthermore, a nine residue stretch of the Δ-domain was shown to be essential for the recruitment of the viral protease (Oh et al., 2014).

Parallels exist between VP1 of SSV1 and gp5 of HK97. It has been shown that the N-terminus of VP1 is essential for infectivity. The N-termini of fusellovirus *vp1* genes are not well conserved with the exception of the essential patch of residues just upstream of the conserved glutamate (Figure 3-7). Complementation of Δ*vp1* mutants with SSV2-*vp1*, which possesses a truncated N-terminus, suggests that the majority of the VP1 N-terminus can be dispensed with as long as the residues upstream of the conserved glutamate remain intact. The role of these essential residues remain mysterious. Similar to the HK97 Δ-

domain, these residues might be critical for recruitment of a protease and/or may act as a scaffold, potentially via interactions with neighboring VP1 N-termini, the minor capsid protein VP3, or some other protein involved in assembly. Research on fusellovirus assembly is very much in its infancy and it is unknown which, if any, fusellovirus proteins acts as a scaffold. Assembly is likely highly conserved among the *Fuselloviridae* and any candidate would likely come from the set of core fusellovirus genes (see chapter 1). The hypothesis that the N-terminus of VP1, at least a portion of which is essential, may fulfill this role is intriguing one and awaits further research.

Conclusions:

We have presented here a comprehensive genetic analysis of the fusellovirus SSV1, which has significantly expanded our knowledge of the genetic requirements for this unique archaeal virus. Almost half of the ORFs encoded by SSV1 were found to be non-essential, while those that are well conserved within the *Fuselloviridae* were essential almost without exception. The near-entirety of the T5 transcript could be deleted without a loss of SSV1 infectivity, raising the question as to why SSV1 devotes a significant portion of its genome to apparently dispensable genes. It is possible these genes were required in the native host but are surplus to requirements in the host organism used here. The effect of such a large mutation on viral fitness remains to be investigated.

SSV1 encodes two apparently unique ORFs, *b49* and *f55*, that are hypothesized to play roles in the response to UV irradiation and the maintenance of the carrier-state (i.e. "lysogeny"), respectively. Both of these genes are apparently not essential for infectivity, but it is not clear what effects these mutations have on the processes they are hypothesized to control. Although little work has focused on B49, a significant amount of research on F55 has recently been conducted and has led to an intriguing hypothesis for the role of F55. Both of these mutants should be incredibly useful for deciphering the roles of these proteins in the SSV1 lifecycle.

Interestingly, infectious viruses harboring mutations within the *vp1* and *vp3* structural genes were able to be isolated. That SSV1 remains infectious after the loss of the minor capsid gene *vp3* is a surprising and unexpected result. We hypothesize that the significant homology that exists between VP1 and VP3 enables the latter to complement this mutant, allowing SSV1 to retain infectivity. Unsurprisingly the *vp1* gene is essential although mutations were tolerated to the universally conserved glutamate residue thought to be involved in the proteolytic maturation of VP1. At least some of the N-terminus of VP1 was also found to be essential for infectivity despite its absence in the mature protein. These results parallel what has been observed in the bacteriophage HK97 and provide insights into the potential role of the N-terminal extension. The impact of the *vp1* and *vp3* mutations on capsid stability and structure remains to be investigated, although

preliminary TEM imaging indicates that Δ*vp3* capsids are elongated relative to the wild-type.

Chapter 4: Purification and characterization of the DNA binding protein VP2 from SSV1

Abstract

Sulfolobus cells possess a number of small and highly-basic chromatin-associated proteins whose biological functions are not entirely understood. Similarly, *Sulfolobus* spindle-shaped virus 1 (SSV1) virions also contain a small, highly-basic DNA binding protein known as VP2. To better characterize the VP2 protein of SSV1, we have developed a protocol for the purification of recombinant 6X His-tagged VP2 from *E. coli*. Purification of 6-His VP2 under denaturing conditions was necessary in order to effectively remove contaminating nucleic acids. However, VP2 could be refolded. 6-His VP2 apparently exists as a monomer in solution and binds non-specifically to dsDNA. A comparison of all known VP2 homologues was performed.

Introduction

Sulfolobus Spindle-shaped virus 1 (SSV1) encodes four structural proteins, VP1, VP2, VP3, and VP4. VP1, VP2, and VP3 were originally identified by N-terminal sequencing of proteins from purified SSV1 virions (Reiter et al., 1987a). VP4 (formerly C792) was identified later via mass spectrometry of purified virions and likely forms the SSV1 tail fiber (Menon et al., 2008; Quemin et al., 2015). VP1 and VP3, the major and minor capsid proteins respectively, are highly conserved among all fuselloviruses (Redder et al., 2009; Iverson and Stedman

2012). VP2 is small (8.83 kDa) and highly basic (pI = 11.39) and is believed to be bound to the viral DNA within the capsid (Reiter, 1985; Reiter et al., 1987a). The genome of SSV1 was reported to exist in a positively supercoiled topological state within the virion (Nadal et al., 1986). It is unclear what role, if any, VP2 plays in the maintenance or establishment of this positive supercoiling.

In contrast to the major and minor capsid proteins, VP2 is much less well conserved among the *Fuselloviridae*. VP2 homologues have been identified in three fusellovirus genomes (SSV1, SSV6, ASV1). Three more VP2-like sequences have been identified in metagenomic analyses from Boiling Springs Lake, an acid mine drainage in Richmond, CA, and a Mexican hot spring (Redder et al., 2009; Andersson and Banfield 2008; Diemer and Stedman unpublished; Servin-Garcidueñas et al., 2013). Because VP2 is poorly conserved and the gene can be deleted from SSV1 without causing an obvious defect, its role is unclear in the fusellovirus life cycle (chapter 2). Perhaps in the absence of VP2 the viral DNA in the virion remains "naked" or a VP2 "substitute" may be recruited from the host.

Members of the genus *Sulfolobus* express several small, basic, nucleoid-associated proteins (NAPs), many of which have been shown to protect the host chromosome from thermal denaturation (Driessen et al., 2011). Two of these proteins, Sul7 and Cren7, comprise roughly 5% and 1% of the total cellular protein content (respectively) and bind non-specifically to DNA as monomers. Both seem to bend DNA to similar degrees by binding non-specifically in the

minor groove of DNA (Driessen et al., 2013). Cren7 and Sul7 exhibit highly similar beta sheet-rich tertiary structures despite the absence of any recognizable sequence homology (Guo et al., 2008). In addition to Sul7 and Cren7, members of the *Sulfolobus* genera also encode a homologue of the highly conserved ALBA (acetylation lowers binding affinity) protein that binds non-specifically to DNA as a dimer (Bell et al., 2002). Binding by Cren7, Sul7 and ALBA all seem to induce negative supercoiling of DNA *in vitro* (Xuan et al., 2012). Another NAP, SMJ12, is found in lower abundances than the other DNA binding proteins (< 0.1% of total cellular protein) and unlike the other NAP's was shown to induce positive supercoiling of DNA mini-circles *in vitro* (Napoli et al., 2001).

As has been detailed elsewhere within this dissertation (see chapters 2 and 3), little is known about the function of the vast majority of fusellovirus proteins. VP2, like most crenarchaeal virus proteins, lacks recognizable sequence homology to proteins of known function. Here we present an optimized purification protocol for recombinant 6X His tagged VP2 and demonstrate its non-specific binding to dsDNA. The results of this work should be useful for future research aimed at elucidating the biochemical and structural properties of this fusellovirus structural protein.

Materials and Methods:

Cloning of the vp2 *gene of SSV1 into pET30 expression vector*

The SSV1 *vp2* gene (NP_039802) was amplified from purified SSV1 DNA (Palm et al., 1991) using Phusion high-fidelity DNA polymerase (Thermo-Fisher). Primers were designed to amplify the entire SSV1 *vp2* gene and included overhangs compatible with the pET30-XA/LIC vector system by Millipore (Table 4-1). PCR was performed with the following cycling conditions: Initial denaturation at 98º C for 3 min; 35 cycles of 98º C for 15 sec, annealing for 15 sec at 50º C, and extension for 15 sec at 72º C; final extension for 3 min at 72º C. Full length PCR-amplified *vp2* DNA was detected on a 1% agarose gel and purified using the PCR Purification Kit according to manufacturer's protocols (Fermentas).

Table 4-1: Primers used for cloning the SSV1 *vp2* gene

Primer	Sequence	Description
VP2_LIC_F [A,B]	**GGTATTGAGGGTCGC**ATGAAG TGGGTGCAAAAGG	Amplifies SSV1_VP2 ORF with LIC overhang
VP2_LIC_R [A,B]	**AGAGGAGAGTTAGAGCC**CTA CTTGCGGTGCATCCG	Amplifies SSV1_VP2 ORF with LIC overhang
pET30_LIPCR_F	ATTGAGGGTCGCATGAAGTG	Removal of N-terminal tag from pET30_VP2
pET30_LIPCR_R	CACCAGACCAGAAGAATG	Removal of N-terminal tag from pET30_VP2
T7 promoter	TAATACGACTCACTATAGGG	Sequencing of pET30_VP2 constructs

[A]Bolded bases denote pET30-XA/LIC LIC overhangs
[B]Underlined bases are SSV1_VP2 start and stop codons

Purified *vp2* DNA was inserted into pET30-XA/LIC expression plasmid (Millipore) and transformed into chemically competent NovaBlue *E. coli* (Millipore) following manufacturer's protocols. Plasmid was purified from a single transformant by alkaline lysis (Birnboim and Doley 1979). Insertion of *vp2* DNA

into pET30-XA/LIC vector was confirmed by EcoRI restriction endonuclease digestion and DNA sequencing (Table 4-1). For expression of recombinant VP2 protein, the pET30_VP2 plasmid was transformed into chemically competent E. cloni BL21(DE3) expression cells following manufacturer's protocol (Lucigen).

Removal of N-terminal tag region from pET30_VP2 expression vector

The recombinant pET30-XA/LIC VP2 construct codes for the 74 amino acids of the native SSV1-VP2 protein with an additional 47 amino acid N-terminus comprised of a 6X-His tag, an S-tag, and factor Xa and thrombin proteolytic cleavage sites (Figure 4-1A). One of the long-term goals of recombinant VP2 expression was to obtain a three-dimensional structure. However, the length of the N-terminal tag could potentially complicate crystallography or NMR experiments. Thus, a large portion of the N-terminus containing most of the tags was deleted using long inverse PCR (LIPCR). Primers were designed to delete 93 bases between the 6X His tag and the start of VP2 gene while maintaining the correct reading frame (Table 4-1). Phusion DNA polymerase was used following manufacturer's protocols with the following PCR cycling conditions: Initial denaturation at 98º C for 3 min; 35 cycles of 98º C for 15 sec, annealing for 15 sec at 50º C, and extension for 2.5 min at 72º C; final extension for 3 min at 72º C. Five identical 20 μL LIPCR reactions were pooled and DNA was purified using the PCR Purification kit (Fermentas) according to manufacturer's protocols. Purified DNA was phosphorylated using T4

polynucleotide kinase following manufacturer's protocols (Fermentas). Phosphorylated DNA was ligated for 20 hours at 16ºC using 5U of T4 DNA ligase (Fermentas) in 20 μL of 1X ligase buffer. 5 μL of the ligation reaction was transformed directly into 100 μL of chemically competent NovaBlue *E. coli* following manufacturer's protocols (Millipore). Plasmid DNA was purified from transformants via alkaline lysis and sequenced to confirm deletion (Table 4-1). This DNA construct is referred to as pET30_6HisVP2 and was utilized to produce 6-His VP2 for all experiments, unless otherwise noted. For expression of 6-His VP2, pET30_6HisVP2 was transformed into chemically competent E. cloni BL21(DE3) expression cells following manufacturer's protocol (Lucigen).

Overexpression of 6-His VP2 protein

Single colonies of E. cloni BL21 DE3 cells harboring pET30_6HisVP2 were inoculated into 5 mL of 2XYT (Sambrook and Russell 2001) medium with 50 μg/mL of kanamycin and grown at 37º C overnight in a shaking incubator (~20 hr). The following morning turbid cultures were diluted 1:100 in fresh 2XYT/Kan medium and allowed to grow for approximately 3 hr until the OD_{600nm} was ~0.40. Cultures were typically 400 mL. IPTG was added to a final concentration of 1 mM and cells were incubated at 37º C with shaking for 4 hr. To check for 6-His VP2 over-expression, 1 mL samples were taken at 0 hr, 2 hr, and 4 hr after the addition of IPTG. Samples were spun in a microcentrifuge at 13,000 x g for 2 min, resuspended in 50 μL of 1X SDS-PAGE sample buffer (see below), heated

at 100ºC for 10 min, and analyzed on 15% SDS-PAGE gels. The OD_{600nm} of 400 mL cultures typically reached values of 1.2-1.4 after 4 hr of incubation. After 4 hr incubation, 100 mL aliquots of overexpressed culture were centrifuged at 3,000 rpm (SA-300 rotor) to pellet cells. Cell pellets were typically stored at -20º C for 1-10 days prior to purification.

Table 4-2: Buffers used in VP2 purification

Buffer Name	Buffer Components
7 M urea lysis buffer	7 M urea, 0.5 M NaCl, 0.1 M NaH_2PO_4, 0.01 M Tris, pH = 8.0
7 M urea wash buffer	7 M urea, 0.5 M NaCl, 0.05 M Na_2HPO_4, 0.01 M Tris, 0.01 M imidazole pH = 8.0
2.5 M urea wash buffer	2.5 M urea, 0.5 M NaCl, 0.05 M Na_2HPO_4, 0.02 M Tris, pH = 8.0
4.5 M urea wash buffer	4.5 M urea, 0.5 M NaCl, 0.05 M Na_2HPO_4, 0.02 M Tris, pH = 8.0
0 M urea wash buffer	7 M urea, 0.5 M NaCl, 0.05 M Na_2HPO_4, 0.02 M Tris, pH = 8.0
Elution buffer	0.1 M NaCl, 0.02 M Tris, 0.5 M Imidazole, pH = 8.0
FPLC equilibration buffer	0.05 M Tris, 0.1 M NaCl, 1% glycerol, pH = 8.0
0.1 M NaCl buffer	0.05 M Tris, 0.1 M NaCl, 1% glycerol, pH = 8.0
2.5 M NaCl buffer	0.05 M Tris, 2.5 M NaCl, 1% glycerol, pH = 8.0
SEC running buffer	0.05 M Tris, 0.1 M NaCl, 5% glycerol, pH = 8.0

Initial purification of 6-His VP2 from E. coli

Cell pellets (from 100 mL cells expressing 6-His VP2) were resuspended in 5 mL of 7 M urea lysis buffer (Table 2) and incubated at room temperature with shaking for one hour. Cells were then sonicated (Misonix Microson) on ice six times for ten seconds each with a two-minute break between pulses. The lysate was spun at 12,000 rpm (SA-300 rotor) for 30 min at room temperature. The VP2-containing supernatant was collected and cell pellets were discarded.

Ni-NTA chromatography

Supernatant from cell lysates containing 6-His VP2 was initially purified using Ni-NTA chromatography. A 10 mL total volume gravity flow column was prepared with a 1 mL bed volume of Ni-nitrilotriacetic agarose (Ni-NTA) resin (Invitrogen) and the column was equilibrated with 10 column volumes (CVs; 10 mL) of 7 M urea lysis buffer (Table 4-2). Approximately 10 mL of 6-His VP2-containing supernatant was loaded onto the column and the flow through was collected. The column was then washed stepwise with 5 CVs of 7 M urea wash buffer, 5 CVs of 4.5 M urea wash buffer, and 5 CVs of 2.5 M urea wash buffer (Table 4-2). A final wash with 10 CVs of 0 M urea wash was performed prior to elution with 5 CVs of elution buffer. 20 μL samples were taken from all eluted fractions and analyzed by SDS-PAGE. Nucleic acid presence in fractions was analyzed by trichloroacetic acid (TCA) precipitation from 100 μL of fractions followed by agarose gel electrophoresis on a 1% gel (see below).

Heparin affinity chromatography

Fractions from Ni-NTA chromatography containing 6-His VP2 were further purified using the AKTA fast protein liquid chromatography platform (FPLC) and Hi-Trap HP 1 mL heparin affinity columns (GE healthcare). The column was equilibrated with 20 CVs of FPLC equilibration buffer (Table 4-2). 10 mL of Ni-NTA fractions containing 6-His VP2 was diluted to 25 mL with FPLC equilibration buffer (Table 4-2) and loaded onto the heparin column. 6-His VP2 was eluted

from the column using a 0.1-2.5 M NaCl gradient in equilibration buffer over 20 CVs. A final wash was performed with 5 CVs of 2.5 M NaCl. All fractions were measured by UV spectroscopy (280 nm) upon elution from the column and 20 μL samples were analyzed by SDS-PAGE. 6-His VP2 typically eluted around 1.0 M NaCl.

Size exclusion chromatography

Fractions containing 6-His VP2 from the Heparin column were further purified via size exclusion chromatography (SEC) using a Superdex 10/300 GL column (GE Healthcare) on an AKTA FPLC platform. The column was calibrated with a Gel Filtration Standard (Bio-Rad #1511901) under the same running conditions. The column was equilibrated in SEC running buffer prior to all experiments (Table 4-2). Heparin-purified 6-His VP2 was concentrated to 100 μL using Vivaspin 20 and Vivaspin 500 centrifugal concentrators (10 kDa MW cutoff) following manufacturer's protocols (Sartorius). 100 μL of concentrated 6-His VP2 was loaded onto the column and fractions were examined via UV spectroscopy (280 nm) upon elution from the column. A 20 μL sample of each fraction was analyzed by SDS-PAGE. 6-His VP2-containing fractions were pooled and the concentration was measured via Bradford assay (Biorad).

Circular dichroism (CD) spectroscopy

SEC-purified 6-His VP2 was submitted to the Bächinger laboratory at OHSU for analysis by far-UV CD spectroscopy to determine if purified 6-His VP2

was folded following purification. CD measurements in millidegrees were performed on 0.1 mg/mL of protein in 20 mM phosphate buffer (pH 7.5) at 25º C using a 0.1 cm path cuvette on an AVIV 202 CD Spectropolarimeter.

Electrophoretic mobility shift assays (EMSAs)

SEC-purified 6-His VP2 was mixed with purified DNA at varying concentrations. DNA used in binding experiments was purified using the GeneJet Plasmid Purification Kit (for plasmid DNA samples) or GeneJet PCR Purification Kit (for PCR amplified DNA samples) following manufacturer's protocols (Thermo-Fisher). Experiments using DNA that was digested with EcoRI were first heat treated at 70º C for 30 min to denature the EcoRI enzyme prior to addition of 6-His VP2. 6-His VP2 and protein were incubated in a 50 mM Tris, 100 mM NaCl (pH = 8) buffer for 30 minutes at room temperature (23ºC) for 30 min prior to loading on a 1% agarose gel. Gels were stained with ethidium bromide and imaged using a UV transilluminator.

Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE)

Samples were prepared for SDS-PAGE analysis as follows. 5X sample buffer (0.25 M Tris-Cl, 0.5 M dithiothreitol, 10% SDS, 0.25% bromophenol blue) was added to a final concentration of 1X and samples were heated at 100º C for 10 minutes. 1-10 μL of prepared sample was loaded onto 15% acrylamide gel in a Tris-glycine buffer system (Sambrook and Russell 2001).

Bradford assay

Protein concentrations were determined using Bradford reagent (Biorad) following manufacturer's protocols using the micro protocol with a bovine serum albumin (BSA) standard.

Trichloroacetic acid (TCA) precipitation

A 1/10 volume of 100% TCA (w/v) was added to the sample in a 1.5 mL tube and mixed by pipetting. Sample was placed at -20º C for 1 hr and centrifuged at 13,000 x g (at 4º C) for 5 min. The supernatant was then removed and discarded without disturbing the small white pellet. The pellet was washed with 200 μL of ice cold acetone (100%) and centrifuged for 5 min at 13,000 x g (at 4º C). The supernatant was removed and the pellet was air-dried for 30 - 60 min. Sample was dissolved in 50 μL of diH_2O and prepared for SDS-PAGE (see above).

Results

Cloning of SSV1 vp2 *and expression of recombinant 6-His VP2*

The *vp2* gene of SSV1 was initially cloned into the pET30-XA/LIC vector (Millipore) for production of recombinant VP2 (Figure 4-1A). In addition to the 6X-His tag, the N-terminus of this vector also encodes an S-tag and factor Xa and thrombin cleavage sites (Figure 4-1B). The cumulative size of the N-terminal tag is nearly equal to the size of the VP2 protein itself, raising concerns that it

could interfere with future crystallography and NMR experiments. To eliminate this as a potential problem, a large portion of the N-terminal tag region was removed by LIPCR to yield the expression vector used for 6-His VP2 protein production (Figure 4-1B). The 6-His VP2 construct used in this work was 90 residues, 74 residues corresponding to the native VP2 protein and 16 residues corresponding to the N-terminal 6X His tag and linker. The theoretical molecular weight of recombinant 6-His VP2 is 10.68 kDa (Figure 4-1B).

A

M*HHHHHH*SSGLVPRGSGMKETAAAKFERNHMDSPELGTGGGSGIEGRMKWV
QKAIKRPGRVHRYLMRLYGKRAFTKDGDIKASYLDKAIKHVKKAKIPKEKKRSLL
SALLLAKRLKRMHRK

B

M*HHHHHH*SSGLVIEGRM**K**WVQ**K**AI**KR**PG**R**VH**R**YLM**R**LYG**KR**AFT**K**DGDI**K**AS
YLD**K**AI**K**HV**KK**A**K**IP**K**E**KKR**SLLSALLLA**KR**L**KR**MH**RK**

Figure 4-1: Amino acid sequences of recombinant VP2 proteins used in this work
(A) Amino acid sequence of full-length recombinant VP2 protein prior to deletion of the tag region. Underlined residues correspond to thrombin site, S-Tag, and Factor Xa site (sequentially from the N-terminus). The amino acids of the native SSV1 VP2 protein are highlighted. The theoretical molecular weight of this protein is13.80 kDa.
(B) Amino acid sequence of recombinant 6-His VP2 protein following removal of N-terminal tags. The amino acids of native SSV1-VP2 are highlighted. Basic residues are in bold face and comprise 26 of the 74 amino acids in native SSV1-VP2 (~35%). The 6-His VP2 protein has a theoretical molecular weight of 10.68 kDa and a predicted pI of 11.35. This protein is referred to as 6-His VP2 and was expressed from the vector pET30_6HisVP2.

Induction of protein production from 400 mL of a transformed *E. coli* culture at OD_{600nm} = 0.4 using 1 mM IPTG consistently produced high yields of recombinant protein four hours post induction (Figure 4-2). No significant increase in protein was observed after longer incubations nor by increasing the amount of IPTG used for induction (data not shown). Although the predicted

molecular weight of recombinant VP2 is 10.68 kDa, the protein migrates at an approximate mass of 15 kDa on SDS-PAGE (Figure 4-2).

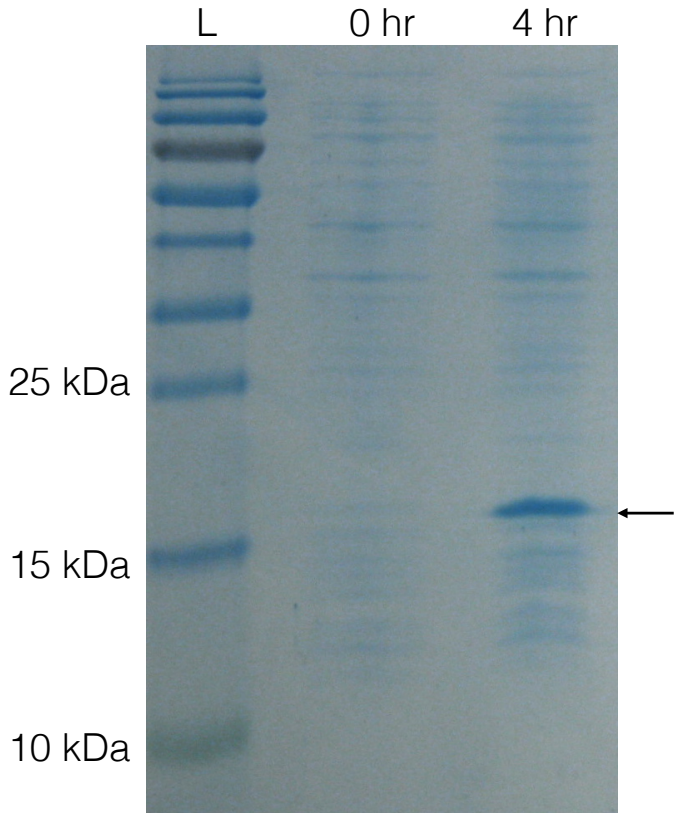


Figure 4-2: Overexpression of 6-His VP2 in BL21 DE3 *E. coli*
A single colony of BL21 DE3 E. cloni (Lucigen) cells harboring the pET30_6HisVP2 expression vector was grown in 400 mL of 2XYT medium to an OD_{600nm} of 0.412. 6-His VP2 production was induced by adding 1 M IPTG to a final concentration of 1 mM. 200 μL samples were taken from the culture just prior to induction (0 hr) and four hours post induction (4 hr). Samples were spun at 13,000 x g, pellets were resuspended in 1X SDS-PAGE sample buffer and analyzed on a 15% SDS-PAGE gel (see methods). The position of 6-His VP2 on the gel is indicated by an arrow. Ladder (L) is PageRuler Prestained Ladder (Fermentas) 10, 15 and 25 kDa proteins are labeled.

Ni-NTA chromatography of recombinant 6-His VP2

Protein preparations were initially purified under non-denaturing conditions. Following protein overexpression, cells were lysed and centrifuged at high speed to separate the cell debris from the supernatant and both fractions were checked for the presence of 6-His VP2 by SDS-PAGE. Although some 6-His VP2 was

always associated with the cell debris, 6-His VP2 was predominantly found in the supernatant (Figure 4-3, lane “PC”). The supernatant was loaded onto a Ni-NTA agarose column to which the 6-His VP2 bound (Figure 4-3A compare PC to FT lanes). Following a wash step, protein was eluted from the column using 500 mM imidazole (Figure 4-3A, lanes E1 to E4). Eluted protein was predominantly 6-His VP2 but small amounts of contaminating proteins were present (Figure 4-3A). Agarose gel electrophoresis of eluted protein fractions (Figure 4-3B) indicated a significant amount of co-purifying nucleic acid that is most likely bound to 6-His VP2 (Figure 4-3B). UV spectroscopy of eluted protein routinely gave 260/280 measurements (1.8 – 1.9) characteristic of nucleic acids.

Because it was later discovered that nucleic acid contamination impaired heparin binding of 6-His VP2 (data not shown), purification was performed under high-salt denaturing conditions (7 M urea and 0.5 M NaCl) to remove contaminating nucleic acids. Cells were resuspended in 7 M urea lysis buffer (Table 4-2), sonicated to lyse cells, and the 6-His VP2-containing supernatant was separated from the cell debris by centrifugation at 12,000 rpm (SA-300 rotor). The supernatant was loaded onto a Ni-NTA agarose column and washed with 7 M urea/0.5 M NaCl buffer (Table 4-2). To re-nature 6-His VP2, the column was washed stepwise with decreasing amounts of urea prior to elution (Figure 4-4A) with a urea-free imidazole buffer (Table 4-2). Agarose gels and spectrophotometry of the eluted fractions indicated that the co-purifying nucleic acid was greatly reduced (data not shown).

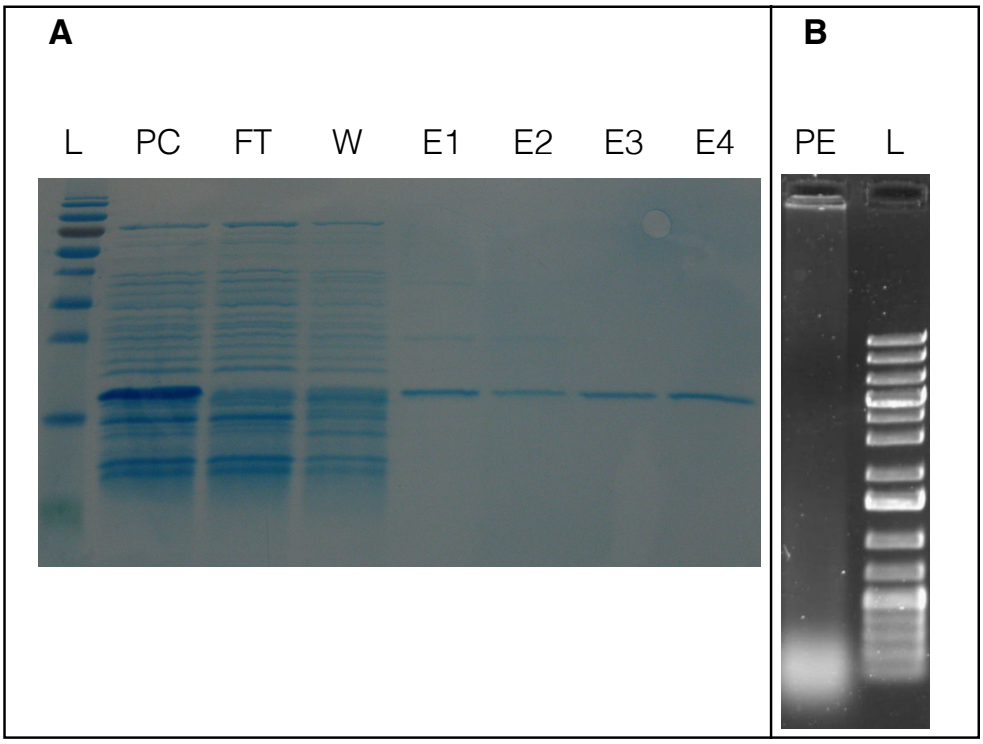


Figure 4-3: Ni-NTA purification of 6-His VP2 under non-denaturing conditions
(**A)** A 400 mL culture of BL21 DE3 E. cloni (Lucigen) harboring pET30_6HisVP2 was treated for 4 hr with 1 mM IPTG, after which the cells were collected by low speed centrifugation (3,000 rpm). Cells were resuspended in 0.5 M NaCl/0.1 M NaH_2PO_4/0.01 M Tris buffer (pH = 8.0) and lysed by sonication (see methods). Lysate was centrifuged at 12,000 rpm (SA-300) and supernatant containing 6-His VP2 was collected (PC). Supernatant was loaded onto Ni-NTA agarose column with a 1 mL bed volume and the flowthrough was collected (FT). The column was washed with 10 column volumes (10 mL) of 0.5 M NaCl, 0.01 M Tris, 20 mM imidazole, pH = 8.0 buffer (W), and eluted with 4 CVs the same buffer containing 500 mM imidazole (E1-E4). All fractions shown on the gel were prepared for SDS-PAGE by diluting 20 μL of sample with 5 μL of 5X sample buffer and heating for 10 min at 100º C. 5 μL of each sample was loaded onto a 15% polyacrylamide gel. Ladder (PR) is PageRuler Prestained Ladder (Fermentas).
(B) Eluted fractions (E1-E4, panel A) from the Ni-NTA column containing 6-His VP2 were pooled and 15 μL was mixed with 3 μL of 6X DNA Loading Buffer (Fisher), separated on a 1% agarose gel, and stained with ethidium bromide (PE). Ladder (L) is GeneRuler 1 kb Plus DNA Ladder (Fisher).

Washing Ni-NTA bound 6-His VP2 with high concentrations of urea (7 M) or NaCl (up to 1.5 M) separately was not as effective as a combination of the two. Nucleic acids were observed in the flowthrough following a wash using only 7 M urea (no NaCl), indicating that urea alone can remove nucleic acid (Figure 4-4B). However, eluted 6-His VP2 containing fractions show that co-purifying nucleic

acids persist by agarose gel electrophoresis (data not shown) and UV spectroscopy (260/280 nm). Likewise, washing Ni-NTA bound 6-His VP2 with 50 CVs of 0.75, 1.0, or 1.5 M NaCl (no urea) was ineffective at total nucleic acid removal after analysis by agarose gels and spectrophotometry (data not shown). Furthermore, washing with high concentrations of NaCl (>1.0 M) resulted in dissociation of 6-His VP2 from the Ni-NTA column.

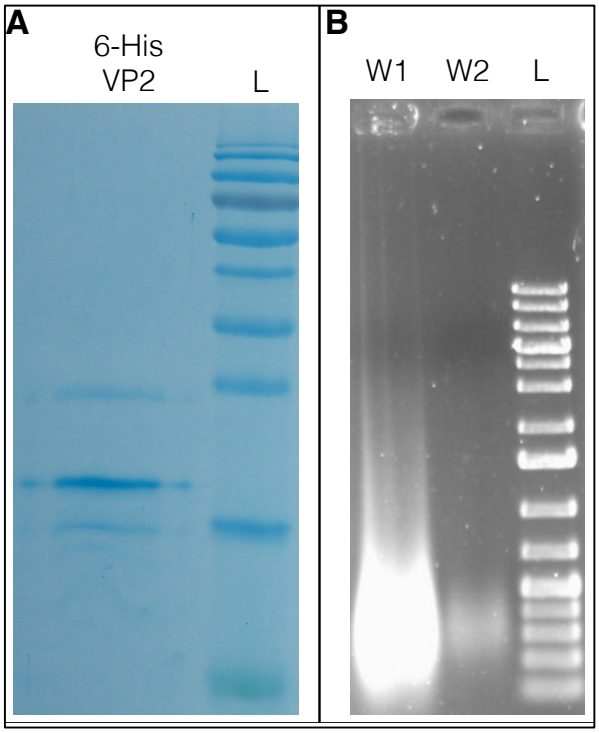


Figure 4-4: Ni-NTA purification of 6-His VP2 using 7 M urea in the absence of NaCl results in incomplete removal of co-purifying nucleic acids
(A) A 400 mL culture of BL21 DE3 E. cloni harboring pET30_6HisVP2 was prepared and protein production was induced as indicated in methods. Following induction, cells were collected by low speed centrifugation (3,000 rpm) and resuspended in 5 mL of 7 M urea buffer (7 M urea, 0.01 M Tris and 0.1 M NaH_2PO_4, pH = 8.0). The cells were sonicated and the lysate was prepared for Ni-NTA chromatography as above (Figure 4-3). Supernatant was loaded onto Ni-NTA agarose column with a 1 mL bed volume and washed with 20 column volumes of 7 M urea buffer. 6-His VP2 was eluted with 5 column volumes of 7 M urea buffer containing 500 mM imidazole. The imidazole elution was pooled and a 20 μL sample was prepared and analyzed by SDS-PAGE as above (6-His VP2). Molecular weight standard (L) is PageRuler Prestained Ladder (Fisher).
(B) The Ni-NTA column with 6-His VP2 bound was washed with 20 column volumes of 7 M urea wash buffer and was collected as two 10 mL fractions (W1 and W2). A 200 μL sample of both fractions was taken and precipitated with TCA (see methods). Samples were loaded onto a 1% agarose gel and stained with ethidium bromide. Ladder (L) is GeneRuler 1 kb Plus (Fisher).

Heparin affinity chromatography

6-His VP2 was further purified by heparin chromatography on an AKTA fast protein liquid chromatography (FPLC) instrument. 6-His VP2-containing fractions from Ni-NTA chromatography were pooled (10-50 mL total) and loaded onto a HiTrap Heparin HP 1 mL column (GE Healthcare) using a superloop. The column was washed with a 2.5 M NaCl gradient over 20 CVs (Figure 4-5A) and 1 mL fractions eluted from the heparin column were analyzed via UV spectroscopy and 15% acrylamide gels (Figure 4-5B). Analysis of eluted 6-His VP2 often resulted in two bands when visualized on SDS-PAGE (see Figure 4-5B), however it is not clear if this is due to degradation of 6-His VP2 or a contaminant. The two bands were only observed at high concentrations of VP2. Alternatively, the second band could be a co-purifying protein of similar molecular weight.

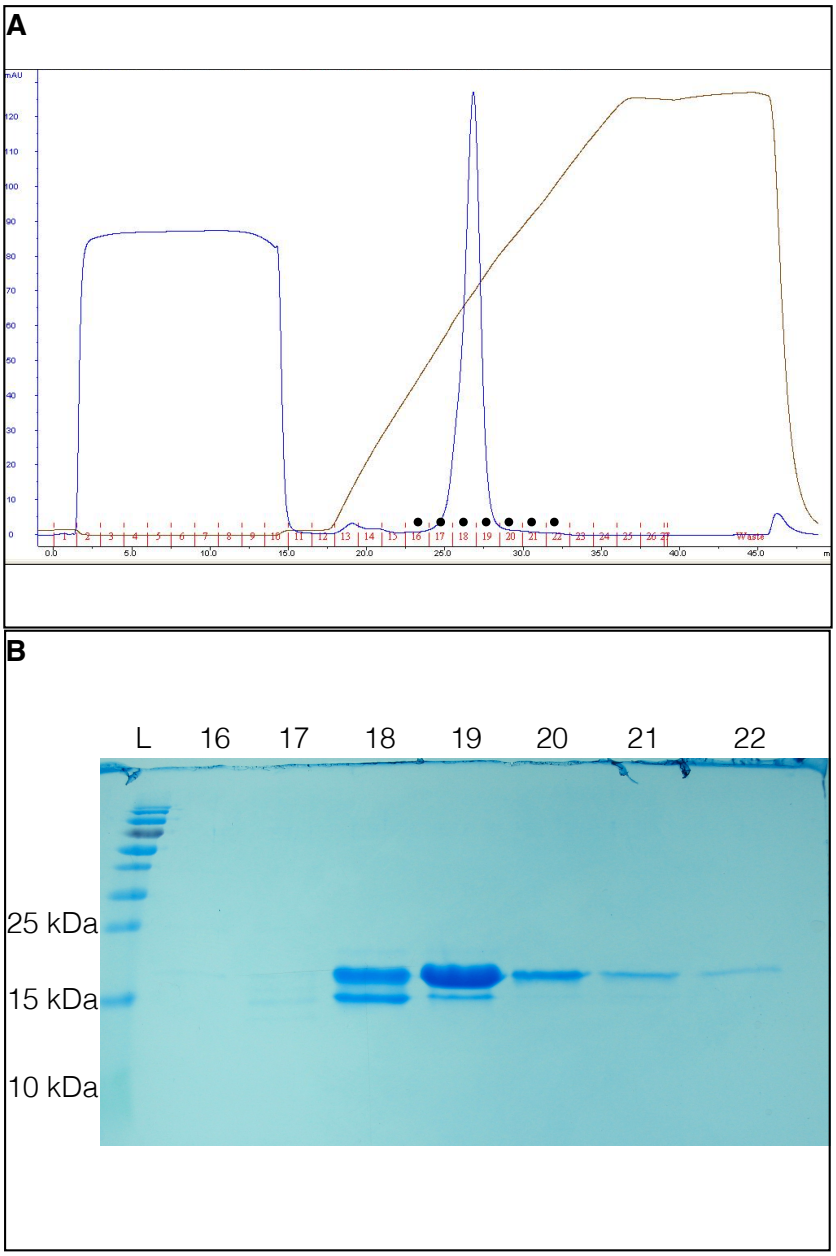


Figure 4-5: Purification of 6-His VP2 by heparin column chromatography
(A) Approximately 10 mL of Ni-NTA purified 6-His VP2 (in elution buffer) was loaded onto a Hi-Trap HP 1 mL heparin affinity column (GE Healthcare) connected to an AKTA fast-protein liquid chromatography system. Bound 6-His VP2 was eluted by washing with a 20 column volume gradient from 0.1 M to 2.5 M NaCl (Table 2). The absorbance at 280nm was monitored (blue line) and 1.5 mL fractions were collected. 6-His VP2 typically eluted around 1 M NaCl. On the chromatogram, the Y-axis is in milli-absorbance units (mAU) and the X-axis is volume (mL). The brown line shows buffer conductivity. Fractions analyzed on the gel in panel B are indicated by filled circles on the chromatogram.
(B) 15% SDS-PAGE gel of fractions following heparin chromatography. 20 μL samples were taken from each fraction and prepared for SDS-PAGE (see methods). The numbers on the top of the gel correspond to fractions from the chromatogram in panel A. Ladder (L) is PageRuler Prestained Ladder (Fermentas).

SEC purification indicates 6-His VP2 exists as a monomer in solution

Following heparin purification, 6-His VP2-containing fractions were pooled, concentrated (~100 μL) and loaded onto a calibrated Superdex 10/300 GL SEC column (GE Healthcare). 6-His VP2 eluted at ~18 mL, corresponding to a molecular mass of ~10 kDa (Figure 4-6) with a "shoulder" eluting later. The theoretical molecular weight of 6-His VP2 is 10.68 kDa, suggesting that 6-His VP2 exists in solution predominantly as a monomer.

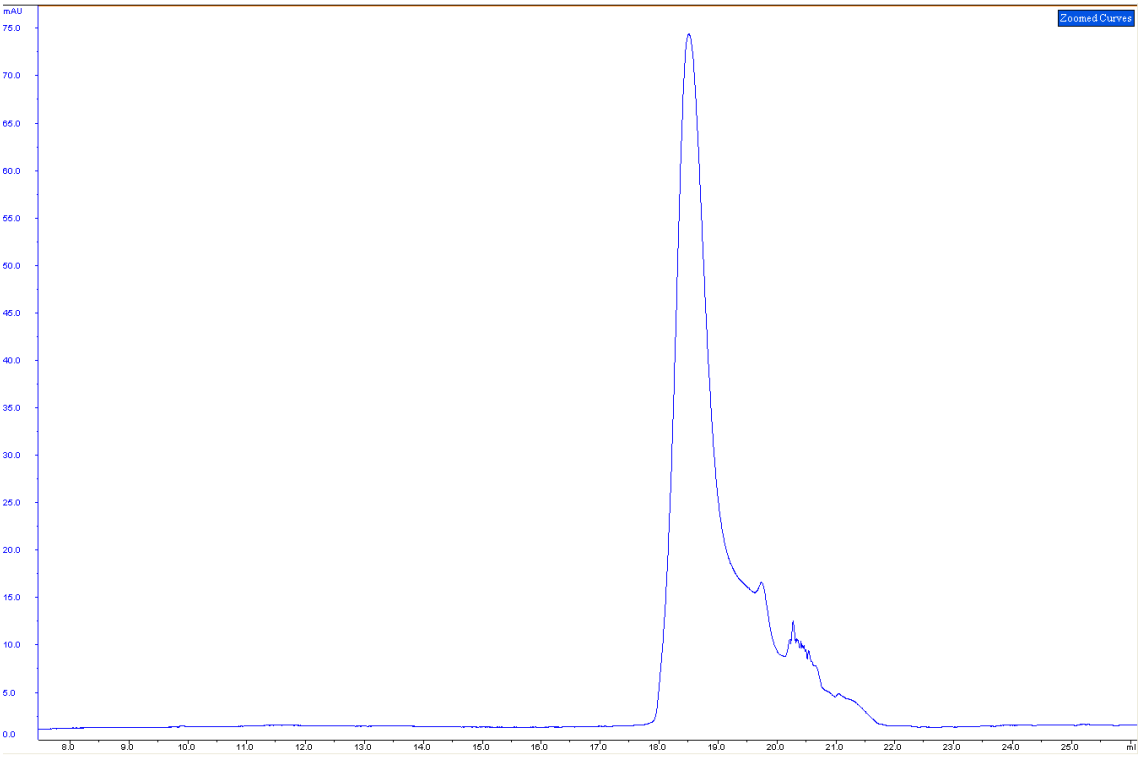


Figure 4-6: Size exclusion chromatography of 6-His VP2
Heparin-purified 6-His VP2 was concentrated to 100 μL (vivaspin 500, Sartorius) and loaded onto a calibrated Superdex 200 10/300 GL column (GE Healthcare) and run at 0.5 mL/min with SEC running buffer (0.05 M Tris, 0.1 M NaCl, 5% glycerol, pH = 8.0). Absorbance at 280 nm was monitored and 1mL fractions (1 mL) were collected. 6-His VP2 eluted at ~18 mL, corresponding to an approximate molecular mass of 10 kDa.

6-His VP2 is folded following purification:

The harsh purification conditions used may have damaged 6-His VP2 and/or led to incorrect folding. Thus it was imperative to confirm that 6-His VP2 had been refolded following the 7 M urea/0.5 M NaCl washes. To check this, circular dichroism (CD) spectroscopy was performed by the Bächinger lab at Oregon Health and Sciences University (Figure 4-7). The spectrum is reminiscent of an α-helix-containing protein, characterized by two broad minima around 208 and 220 nm and a maximum at ~190 nm (Figure 4-7, arrows). These results indicate that the protein is predominantly alpha helical and imply that 6-His VP2 is in its native folded state. Furthermore, a heparin-column purified fraction of 6-His VP2 (1.3 mg/mL) was analyzed by collaborators at Montana State University using NMR. The chemical shifts in the NMR spectrum were reportedly well resolved, confirming that the protein is in fact folded (data not shown).

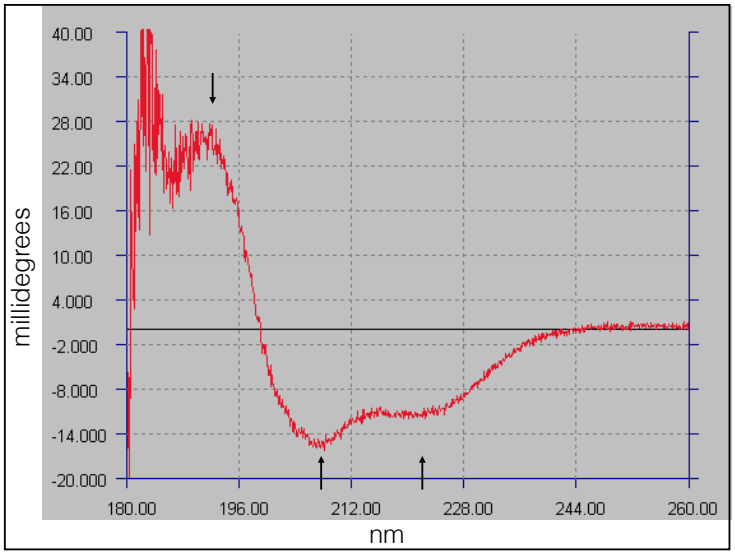


Figure 4-7: CD spectrum of 6-His VP2 at 25º C
SEC-purified 6-His VP2 was diluted to 0.1 mg/mL in 20 mM phosphate buffer (pH = 7.5) and analyzed by far-UV CD spectroscopy using a 0.1 cm path-length cuvette. Arrows denote the two characteristic absorbance minima (208 and 220 nm) and the absorbance maximum (190 nm) typical of α-helix-containing proteins. Y-axis units are in millidegrees and X-axis units are in nanometers (nm).

DNA binding by 6-His VP2:

VP2 purified from SSV1 virions was previously shown to bind non-specifically to dsDNA (Reiter 1985). Similarly, SEC-purified 6-His VP2 appears to bind both linear and circular dsDNA (Figure 4-8). Increasing amounts of 6-His VP2 were combined with undigested pAJC97 DNA (a SSV1 shuttle vector containing the complete SSV1 genome and an *E. coli* plasmid, Table 3-3) and incubated at room temperature for 30 minutes prior to loading onto a 1% agarose gel (Figure 4-8A). The DNA begins to shift at a protein:DNA molar ratio of 585. At a protein:DNA molar ratio of 5850 the DNA appears to be completely retained in the well of the agarose gel. At higher protein:DNA ratios the DNA is no longer visible on the gel (Figure 4-8A). An almost identical binding pattern was observed when pAJC97 DNA was digested with EcoRI prior to VP2 binding (Figure 4-8B). Increasing the incubation time to two hours prior to loading onto the gel had no effect and did not alter the results (data not shown).

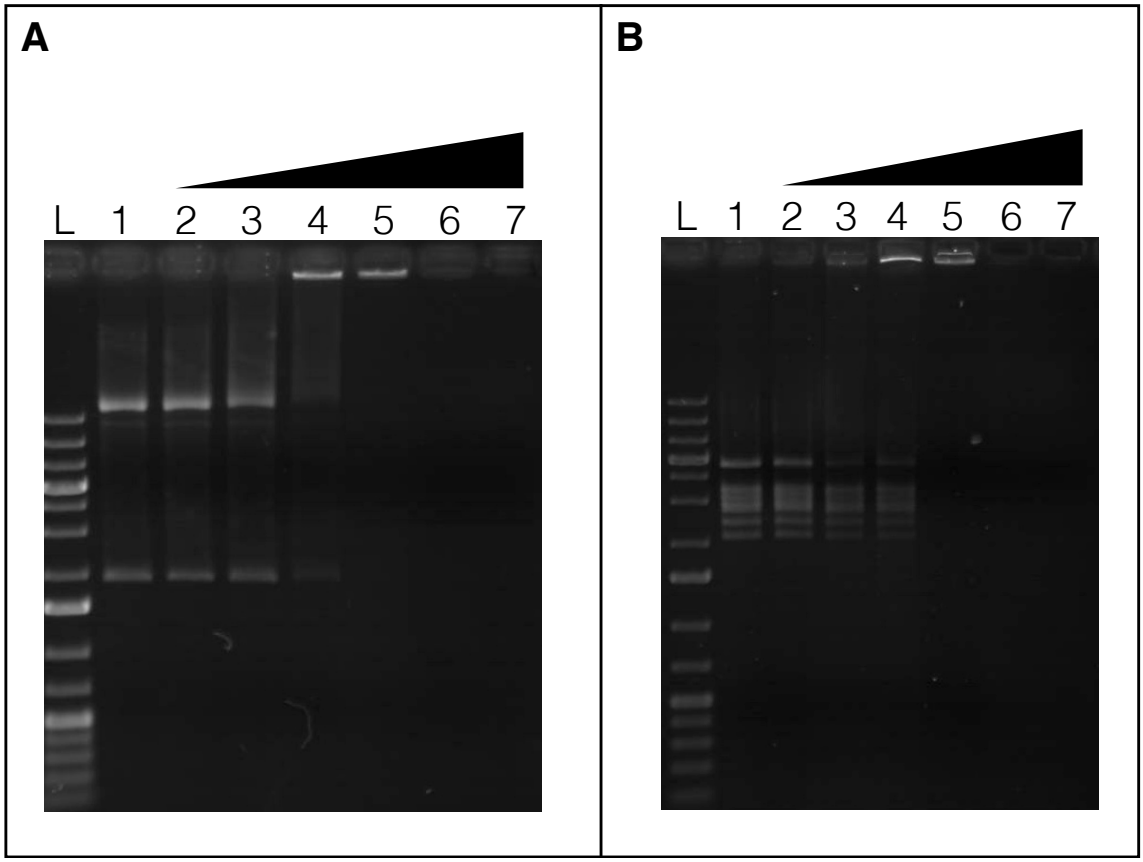


Figure 4-8: 6-His VP2 binding of circular and EcoRI-digested pAJC97 (SSV1) DNA
pAJC97 DNA (A – undigested pAJC97; B – EcoRI-digested pAJC97) and SEC-purified 6-His VP2 were mixed together in varying ratios (see below), incubated at room temperature for 30 min, and run for 45 minutes on a 1% agarose gel. Gels were stained in ethidium bromide and imaged using a UV transilluminator. The numbers on the top of each gel correspond to the following amount of VP2: 1 – 0 μg VP2; 2 – 0.05 μg; 3 – 0.1 μg; 4 – 0.5 μg; 5 – 1 μg; 6 – 5 μg; 7 – 10 μg. The total amount of DNA (undigested and digested) in each lane is 200 ng. Ladder (L) is GeneRuler 1 kb plus (Fermentas).

Discussion and conclusions

The role of VP2 in the SSV1 life cycle is poorly understood. Apart from its presence in the SSV1 virion, little is known about the three-dimensional structure or DNA binding affinity and specificity of VP2. We demonstrate here an effective protocol for the expression and purification of recombinant 6X His-tagged VP2 from *E. coli*. The results of this work should prove useful for future research aimed at characterizing VP2 and elucidating its physiological role.

Purification of recombinant 6-His VP2

Initially, the VP2 gene from SSV1 was cloned into the pET30 XA/LIC vector and recombinant protein was purified as detailed above. Concerned over the possibility that the 47-residue N-terminal tag could interfere with future structural experiments, we decided to remove 31 residues from the expression vector using LIPCR (Figure 1). This region could, in theory, have been removed post-expression using Factor Xa protease (Millipore) but we decided against this extra step and possible necessity of another purification step to simply remove it from the vector itself.

Purification of 6-His VP2 was performed under denaturing conditions with 7 M urea in order to facilitate removal of co-purifying nucleic acids following overexpression in *E. coli*. *E. coli* nucleic acid was most efficiently removed by thoroughly washing Ni-NTA-bound 6-His VP2 with 7 M urea and 0.5 M NaCl together. Washing with 7 M urea and 0.5 M NaCl individually did not effectively remove contaminating nucleic acid. Higher concentrations of NaCl were also explored (up to 1.5 M) but were similarly unsuccessful at removing nucleic acid and had the added drawback of causing 6-His VP2 to prematurely dissociate from the Ni-NTA column. That VP2 apparently binds well to *E. coli* DNA is unsurprising given that VP2 was reported in a thesis and review articles to be a DNA binding protein (Reiter 1985; Reiter et al., 1987a; Wiedenheft et al., 2004). It is also not unheard of for *Sulfolobus* virus DNA binding proteins to be contaminated with difficult to remove nucleic acids during the purification process

(Erdmann et al., 2014a). Prior to elution from the column, 6-His VP2 was washed with decreasing concentrations of urea to re-nature the protein. CD spectroscopy was performed on SEC-purified protein and indicated the presence of α-helical regions (Figure 4-7), suggesting that 6-His VP2 was folded. Preliminary NMR experiments by our collaborators at Montana State University also indicated SEC-purified 6-His VP2 has defined structure and further suggest that the protein is not irreversibly damaged during the harsh purification process.

6-His VP2 exists as a monomer in solution

Purification of 6-His VP2 via size exclusion chromatography suggests that 6-His VP2 exists in solution as a monomer, although it may oligomerize upon binding to DNA (Figure 4-6 and Figure 4-8). Monomeric nucleoid-associated proteins (NAPs) are not unusual in *Sulfolobus* which encode several small and highly basic proteins that bind DNA non-specifically (Driessen et al., 2013). A multiple sequence alignment of VP2 with monomeric NAPs Cren7 and Sso7d from *S. solfataricus* P2 revealed no detectable homology (data not shown). However, the protein sequences of Cren7 and Sso7d also do not share any homology despite having nearly identical 3D structures (Guo et al., 2008). Therefore, it is not unreasonable to speculate that VP2 may possess a similar structure despite the absence of homology. It's clear that structural studies are required to better understand the exact role of VP2 and its relationship to other NAPs in *Sulfolobus*.

The presence of α-helical secondary structure in VP2 is not unexpected as it was shown that 6-His VP2 can bind DNA (Figure 4-8). Helix-turn-helix (HTH) DNA binding motifs are comprised of α-helices and are not uncommon in crenarchaeal virus proteins, including the products of SSV1 ORFs *b115, f112*, and *f93* (Prangishvili et al., 2006; Menon et al., 2008; Kraft et al., 2004a). Furthermore, a secondary structure prediction (JPred4) with the native VP2 sequence predicts several α-helical segments (Drozdetskiy et al., 2015; data not shown). VP2 may potentially encode HTH-like domain, however, this awaits structural characterization.

6-His VP2 binds non-specifically to linear and circular dsDNA

Purified VP2 was shown to bind both linear and circular dsDNA and this binding appears indistinguishable between the two forms of DNA (Figure 4-8). These data strongly suggest that VP2 is indeed a DNA binding protein as previously hypothesized (Reiter 1985) but do not conclusively show that VP2 binding is non-specific as binding to non-SSV1 DNA was not demonstrated. Because sequence-specific DNA binding proteins are capable of binding non-specifically to DNA at high protein concentrations, it cannot be entirely ruled out that VP2 does not recognize a specific sequence (Menon et al., 2008). However, the experimental evidence presented here along with the hypothesized role of VP2 does not support this.

Comparison of VP2 homologues

VP2 homologues have been identified in three fusellovirus genomes (SSV1, SSV6, and ASV1) and three environmental metagenomes (Redder et al., 2009; Andersson and Banfield 2008; Servin-Garcidueñas et al., 2013; Diemer and Stedman unpublished). All VP2 homologues are small (< 14 kDa) and have high predicted isoelectric points, similar to other *Sulfolobus* NAPs (Table 4-3). A multiple sequence alignment illustrates the high degree of conservation among all fusellovirus VP2 homologues (Figure 4-9A). Redder et al. stated that the C-termini of VP2 differs among known homologues and that this may reflect alternative modes of interaction between VP2 and the capsid proteins VP1 and/or VP3 (Redder et al., 2009). Although a small C-terminal region is hyper-variable among the different fusellovirus VP2 homologues, the majority of the C-termini appear to be highly conserved (Figure 4-9A). The exception appears to be SMF1-VP2 which differs from the other VP2 homologues by the presence of a short C-terminal extension that is rich in histidine residues. It is worth noting, however, that the genome of SMF1 was assembled from a metagenome and no virion was isolated (Servin-Garcidueñas et al., 2013). Additionally, the SMF1 genome apparently lacks two of the genes belonging to the so-called "fusellovirus core" (i.e. genes conserved in all fuselloviruses; see Table 1-2), casting doubt onto whether or not a SMF1 virion truly exists (Appendix B).

Table 4-3: Properties of VP2 homologues and *S. solfataricus* P2 NAPs

Protein	Molecular Weight (kDa)	Length (amino acids)	Isoelectric point (pI)	Accession
SSV1_VP2	8.8	74	11.39	NP_039802
SSV6_VP2	9.16	76	10.98	YP_003331451
ASV1_VP2	10	82	10.9	YP_003331406
SMF1_VP2	10.44	85	10.56	YP_007678026
AMD_VP2 [A]	13.52	115	9.85	ACD75427
BSL_VP2 [B]	5.68	49	10.24	n/a
Sso7d [C]	7.7	68	9.52	NP_343889.1
Cren7 [C]	6.6	60	9.87	NP_342459.1
ALBA [C]	10.9	100	10.33	NP_342446.1

[A] Andersson and Banfield 2008
[B] Diemer and Stedman unpublished data
[C] *S. solfataricus* P2 genome

A multiple sequence alignment comparing the fusellovirus VP2 homologues and two putative VP2 homologues identified via metagenomic surveys was performed (Figure 4-9B). VP2-AMD, a putative VP2 homologue from an acid-mine drainage metagenome (Andersson and Banfield 2008), is significantly larger than the other VP2 homologues due to the presence of a non-conserved N-terminal extension (~30 residues). This extension, similar to the entire protein, is highly basic (pI ~ 10.9) and is presumably involved in nucleic acid binding. The C-terminus of AMD-VP2 is poorly conserved, suggesting that this region may be involved in different interactions compared to the other fusellovirus VP2 homologues. The putative VP2 homologue identified in the Boiling Springs Lake metagenome (BSL-VP2; Diemer and Stedman unpublished) shows clear homology with the N-termini of the fusellovirus VP2 sequences (Figure 4-9B). However, the sequence is truncated and may not represent the complete protein sequence.

The identification of any fusellovirus genes in an acid mine drainage sample is intriguing as it is the only VP2 homologue found in a non-thermal environment (39º C) and expands the potential habitats from which fuselloviruses could be isolated (Andersson and Banfield 2008). Although only six VP2 homologues have been identified, those that have been all come from hyper-acidic environments, supporting the hypothesis that VP2 helps protect viral DNA from the harsh conditions (Reiter et al., 1987a; Reiter 1985). It is curious that none of the better conserved fusellovirus "core" genes (See Table 1-2) were identified along with VP2 in either the acid mine drainage or the BSL metagenomes (Andersson and Banfield 2008; Diemer and Stedman unpublished). While this could be attributed to the stochacisty associated with metagenomic sampling, it is entirely possible that these VP2 homologues are carried by different viruses (or alternatively, cells) which could expand our knowledge of horizontal gene transfer among viruses in extreme environments.

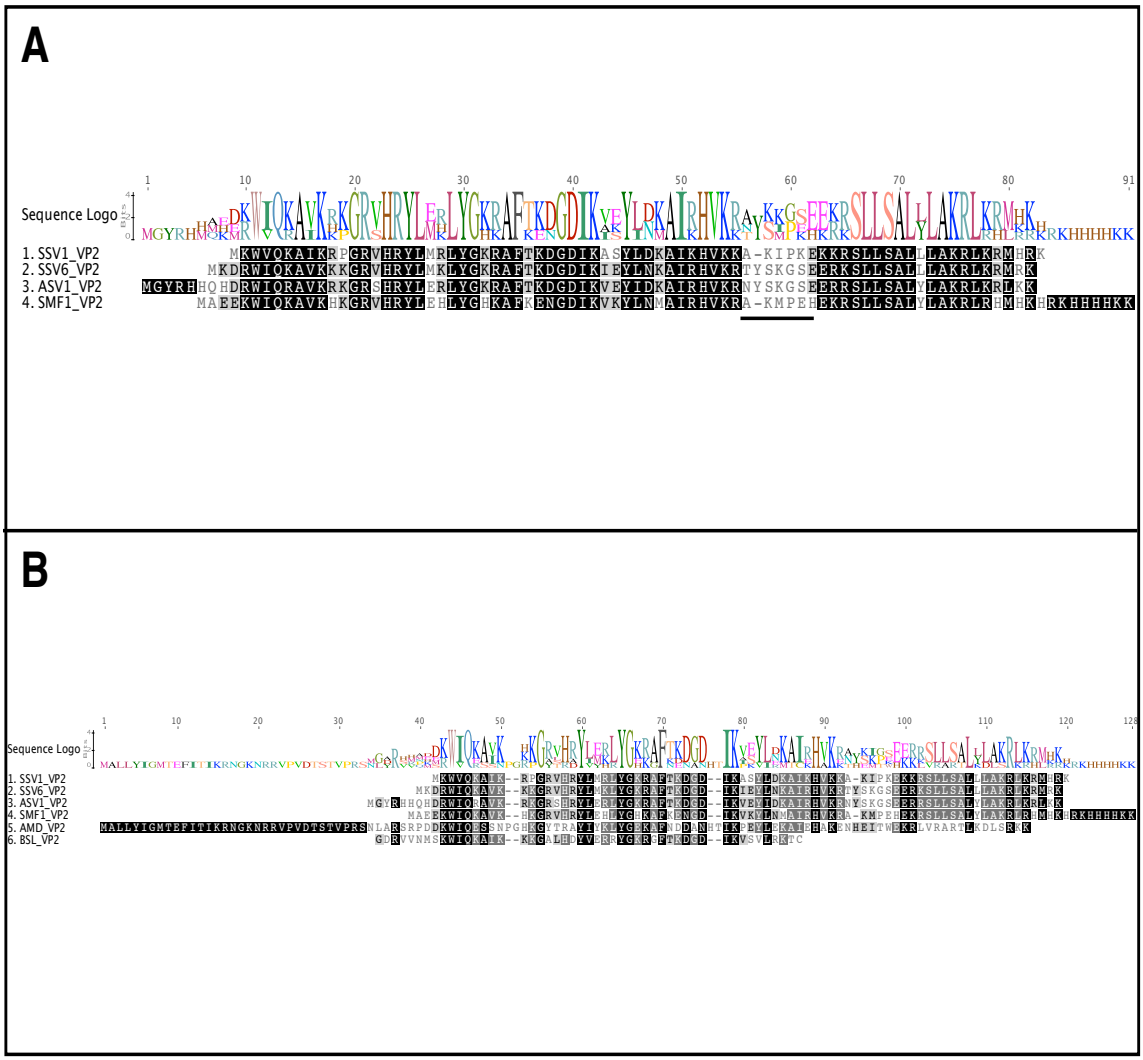


Figure 4-9: Multiple sequence alignments of VP2 homologues
A global alignment of the four fusellovirus VP2 homologues (Panel A) and all known VP2 homologues (Panel B) was performed using Geneious software with default parameters (Biomatters). A sequence logo was generated for both alignments and indicates the degree of conservation at each position in the alignment. Completely conserved residues are highlighted in black. The underlined region in panel A indicates the C-terminal hyper-variable region previously observed in fusellovirus VP2 homologues (Redder et al., 2009).

Putative role of VP2 in the Fuselloviridae

The role of VP2 in the fusellovirus lifecycle is unclear, as the gene is poorly conserved and apparently non-essential for SSV1 infectivity (Chapter 2). A link

between the positively supercoiled state of the viral genome and VP2 might exist, however it remains to be seen if fusellovirus virions lacking VP2 homologues contain positively supercoiled DNA. The observation that reverse gyrase can induce positive supercoils *in vitro* and the fact that it is present in almost every hyperthermophilic Archaea and Bacteria led to the hypothesis that positively supercoiled DNA is more stable at high temperatures (Brochier-Armanet and Forterre 2007). This has been challenged by experiments showing positively supercoiled DNA is no more stable than negatively supercoiled DNA at high temperatures as well as the observation of a number of plasmids in thermophiles that exist in relaxed and negatively supercoiled topologies (Marguet and Forterre 1994; Lulchev and Klostermeier 2014). The almost universal conservation of the reverse gyrase gene in bacterial and archaeal thermophiles strongly suggests a role in the adaptation to life in thermophilic environments, however, the specifics of this role appear complex and remain obscure (Brochier-Armanet and Forterre 2007).

Because it is not clear if the positively supercoiled genome of SSV1 arises from the action of reverse gyrase, it is fair to speculate that VP2 may be responsible. The NAP SMJ12 from *S. solfataricus* P2 was reported to induce positive supercoiling of DNA mini circles *in vitro* and VP2 may perform a similar function in SSV1 (Napoli et al., 2001). Alternatively, VP2 could play a role in binding and stabilizing positively supercoiled DNA that is created by reverse gyrase or some other enzyme. While the action of VP2 is clearly not crucial for

virus function, its presence may confer an advantage to viruses encoding a VP2 homolog.

If VP2 does confer an advantage, then why don't most other fuselloviruses encode such a protein when they inhabit analogous niches? It is not unreasonable to suppose that other fuselloviruses have simply dispensed with encoding their own VP2 homologues and have instead co-opted a host DNA binding protein for the role. Viral genomes are typically streamlined to encode only the minimum information necessary to function, thus the loss of an un-needed gene could be advantageous (Cann 2016). In support of this hypothesis, purified SSV1 virions were found to contain significant amounts of the host-derived NAP Sso7d in addition to VP2 (Quemin et al., 2015).

Purified STIV virions have also been found to contain Sso7d (Maaty et al., 2006). Cells infected with STIV have also been shown to up-regulate expression of Sso7d genes following infection with STIV (Ortmann et al., 2008). There is no evidence for STIV having possessed and subsequently lost its own VP2-like protein, but this is a clear example of a crenarchaeal virus containing a host-derived DNA binding protein. Because this has only been observed in STIV and SSV1, it seems likely that this phenomenon is wide-spread among the crenarchaeal viruses. The structural protein profiles of a number of crenarchaeal virions have been analyzed but none appear to include obvious viral or host derived DNA binding proteins (Prangishvili et al., 2006; Xiang et al., 2005; Geslin et al., 2007; Erdmann et al., 2014b). This includes all isolated hyperthermophilic

spindle-shaped viruses, none of which encode a homologue of the VP2 gene (i.e. ATV, STSV1, STSV2, APSV1, and PAV1).

Chapter 5: First insights into the kinetics of the fusellovirus replication cycle

Abstract

Archaeal virus life-cycles and virus-host interactions are poorly understood. To begin to address this we have initiated one-step growth curve experiments using the well-characterized fusellovirus *Sulfolobus* spindle-shaped virus 1 (SSV1). A "one-step growth curve" has not been performed on SSV1 or any fusellovirus due to difficulties in obtaining the large amount of virus required. We show here that infection of *S. solfataricus* with SSV1 and SSV1 mutants at a low MOI results in the production of viral titers in excess of 10^9 PFU/mL. Virus from low MOI infections was isolated and used to conduct infections at an MOI of 5. Infection with SSV1 at a high MOI results in a severe growth defect that is reminiscent to infection with the related fusellovirus SSV9 at a low MOI. An SSV1 mutant lacking the *d244* ORF elicited a dramatic cellular response following infection at high MOI and produces a viral burst size that is 10-fold lower than wild type SSV1. We also show that UV-inactivated SSV1 does not result in any detectable growth defect in *S. solfataricus*, contrasting what has been reported for SSV9 in S. *islandicus*.

Introduction

The viruses infecting Archaea appear to be extremely diverse both in terms of their morphology and genomic content. Archaeal virus genomes

encode few sequences with recognizable homologues in the public databases (Prangishvili et al., 2006). This dearth of information regarding protein function confounds our understanding of archaeal viruses. Spindle-shaped viruses, which have only been found to infect archaeal hosts, appear widespread across a variety of habitats (Krupovic et al., 2014). The spindle-shaped viruses of the family *Fuselloviridae* are among the best-characterized archaeal viruses, however, very little is known about fusellovirus life cycles and the relationships between these viruses and their hosts.

One-step growth curves have been a cornerstone of virology research since the pioneering work by Max Delbruck and colleagues in the early 20th century (Ellis and Delbruck 1939). They illustrated that a typical virus infection can be divided into three phases: adsorption, replication, and lysis. The crux of the one-step growth curve is synchronous infection, whereby all cells in a given culture are infected by at least one virus at the same time. Synchronous infections require high viral titers (~10^9 – 10^{10} PFU/mL) which have been difficult to attain for *Sulfolobus* spindle-shaped virus 1 (SSV1). Therefore, one-step growth curves have not yet been performed with SSV1, which limits our knowledge of the SSV1 life cycle.

Sulfolobus spindle-shaped virus 1 (SSV1) was isolated from a geothermal hot spring in Beppu, Japan and is one of the best-characterized archaeal viruses to date (Yeats et al., 1982; Prangishvili 2013). The 15.465 kb circular dsDNA genome of SSV1 encodes 35 open reading frames (ORFs), of which only five

have been assigned functions (Iverson and Stedman 2012; Quemin et al., 2015). Clues into the functions of some of the remaining proteins have been provided by structural studies, bioinformatics, and mutagenesis, although the activity of the majority remain elusive (Prangishvili et al., 2006; Dellas et al., 2013; Chapters 2 and 3). Deletion of the ORF *d244*, which encodes a putative nuclease, did not impair SSV1 infectivity, however, cells infected with virus lacking ORF *d244* displayed a retarded growth phenotype compared to cells infected with wild-type virus (Chapters 2 and 3).

Following infection, the SSV1 genome integrates site-specifically into an arginyl tRNA gene, resulting in splitting of the apparently non-essential viral integrase gene (Reiter et al., 1987; Schleper et al., 1992; Clore and Stedman 2006; Chapter 3). SSV1 establishes a stable carrier-state within the infected cell and infected cells often show little-to-no impairment of growth (Fusco et al., 2013; Schleper et al., 1992; Chapter 2). During this carrier-state, the SSV1 genome is maintained at a low copy number and only a subset of viral genes are expressed (Fusco et al., 2013; Fusco et al., 2015c). SSV1 infection does not appear to result in lysis, and viral release probably occurs via budding (Martin et al., 1984; Quemin et al., 2015).

SSV1 is the only fusellovirus whose replication has been shown to be UV-inducible (Martin et al., 1984). Following exposure to UV light, the titer of SSV1 reached ~10^8 PFU/mL within 8 hours of UV irradiation (Schleper et al., 1992). UV-irradiated infected cells subsequently displayed a dramatic decrease in

growth rate compared to the uninfected UV-irradiated control (Schleper et al., 1992). Similar SSV1 titers were achieved in a more recent study using similar UV-induction (Fusco et al., 2015a). The transcription cycle of SSV1 in a stably-infected cell following exposure to UV light was analyzed and shown to be temporally regulated (Fröls et al., 2007). The transcription cycle following an unambiguously synchronous infection by SSV1, or any fusellovirus, remains to be analyzed.

Several other fuselloviruses have been isolated and characterized, though not much has been reported about the interactions with their respective hosts (Ceballos et al., 2013; Prangishvili 2013). SSV2 and SSV9 are the most studied fuselloviruses other than SSV1. SSV2 was originally isolated from a culture of *S. islandicus* strain REY15/4 but has been shown to also infect *S. solfataricus* strains (Arnold et al., 1999; Ceballos et al., 2013). SSV1 and SSV2 are very similar in gene content (21 homologous open reading frames) and gene synteny (Stedman et al., 2003; Appendix B). The transcription cycle of SSV2 was analyzed following infection at an unreported MOI and found to be temporally regulated, similar to SSV1 (Ren et al., 2013). SSV2 replication is induced when the *S. islandicus* REY15/4 culture reaches late logarithmic or stationary growth (Contursi et al., 2006). Interestingly, induction of replication was not observed in *S. solfataricus* strains infected with SSV2 (Contursi et al., 2006). Recently it was shown that infection of *S. solfataricus* strain LnF1 (uracil auxotroph of *S. solfataricus* P2) with SSV2 at an unknown multiplicity of infection (MOI) elicits a

strong response from the host, including activation of the CRISPR-Cas system, ultimately resulting in a reduction of the SSV2 copy number (Fusco et al., 2015c). However, infection of the same host with SSV1 (also at unknown MOI) apparently did not result in CRISPR-Cas activation nor did a co-infection with both SSV1 and SSV2 (Fusco et al., 2015c).

SSV9 (formerly *Sulfolobus* spindle-shaped virus Kamchatka 1) was isolated from the Valley of the Geysers in the Kamchtakan peninsula in eastern Russia (Wiedenheft et al., 2004). Recently it was shown that incubation of *S. islandicus* RJW002 with SSV9 preparations at an MOI of 0.01 resulted in a population-wide dormant-like state, suggesting that very low viral abundances can cause dramatic population-wide effects (Bautista et al., 2015). They also showed that SSV9 was unable to replicate in these cells and suggest their data support the hypothesis put forth by Makarova et al. that dormancy may be a natural reaction to allow the cells time to mount an immune response (via the CRISPR-Cas pathway) and eliminate the infection (Makarova et al., 2012). Interestingly, cell dormancy was also triggered by UV-inactivated virus (i.e. non-infectious) particles, indicating that the mere presence of virus is enough to cause cells to enter stasis.

The unrelated lytic crenarchaeal viruses STIV1 and SIRV2 have also been relatively well studied. Cells infected with SIRV2 at a MOI of 7 cease growing within 4 hours post infection (h.p.i.) and remain in this state of arrest for 60 hr, after which there is a burst of growth from a small subset of the population that is

immune to SIRV-2 infection (Bize et al., 2009). A one-step growth curve of an SIRV2 infection of *S. solfataricus* 5E6 cells at a MOI of 30 caused a similar reduction in host growth to the previous work and revealed a ~5 hr virus latent period with a large burst in viral titer (to $>10^9$ PFU/mL) at 12 h.p.i. (Okutan et al., 2013). Infection of a *S. solfataricus* P2 derivative with STIV1 at an MOI of 1.5 – 2 resulted in a reduction in cell growth at ~30 h.p.i. The copy number of STIV peaked at ~40 h.p.i. (Ortmann et al., 2008). Most of the host cells appeared lysed by 40 h.p.i. (Brumfield et al., 2009).

To better understand the fusellovirus infection cycle we have initiated one-step growth curve experiments on SSV1 and SSV1 mutants. In this work we show that infection of *S. solfataricus* S441 cells with SSV1 at a low MOI results in a burst in virus production at about 48 h.p.i. with titers reaching ~10^9 PFU/mL. The production of high viral titers allowed isolation of sufficient infectious virions to perform higher MOI infections. High MOI infections resulted in a dramatically reduced culture growth rate and resulted in the production of infectious virus by ~12 h.p.i. Growth curves from an SSV1 mutant harboring a deletion of the *d244* ORF display aberrant cell growth and virus production relative to infection with other viruses. We also show that infection of S441 with SSV9 at a low MOI results in an abrupt and significant growth defect. Finally, non-infectious UV-inactivated SSV1 apparently elicits no response from the host, contrary to what has been reported for SSV9.

Materials and Methods:

Growth of Sulfolobus cultures

All *Sulfolobus* cultures were grown in Yeast-Sucrose (YS) medium at pH 3.0 in a 75º C shaking incubator (See chapter 2). Typically, small amounts of frozen cell stocks (ca. 50 μL) were inoculated into 5 mL of YS medium and incubated until turbid (48 – 96 hr). Cells were then transferred to a larger volume of YS medium (50 – 100 mL) in long-neck Erlenmeyer flasks and incubated until the desired OD_{600nm} was reached.

Comparison of filtered and non-filtered virus supernatants

SSV1-infected (S524) and SSV1(*e178::Topo*)-infected (S538) cultures (see below and table 5-1) were grown from frozen stocks in 50 mL of YS medium at 75º C for 5 days. The cell density (OD_{600nm}) of S524 was 0.854 and of S538 was 0.655 when cultures were removed from high temperature. Cultures were centrifuged for 10 min at 6,000 x g (Eppendorf 5810R) at room temperature. The supernatant was collected and split into two aliquots of equal volume. One aliquot (~20 mL) of each sample was passed through a 0.45 μm filter (Sartorious minisart SFCA) and collected. The other aliquot (~20 mL) was not filtered. The titer of the filtered and unfiltered aliquot from both samples was measured by plaque assay in quadruplicate.

Initial purification of virus for one-step growth curves

S. solfataricus S441 was initially transfected with SSV1 by electroporation using purified SSV1 DNA as detailed previously (Chapter 3). Electroporated cells were grown in 50 mL of YS medium and incubated for 120 hours. The cell density (OD_{600nm}) and viral titer were measured every 24 hours. The average viral titer was determined from plaque assays performed in triplicate on cell-free supernatants. Glycerol stocks were prepared from this infected culture (strain S524) and used to propagate SSV1 for future experiments (see chapter 3 methods). A similar protocol was used to prepare frozen stocks of cultures stably infected with SSV1(e178::Topo) and SSV1*Δd244*(*e178::Topo*) (S538 and S530, respectively; see Table 5-1).

SSV1 production in stably-infected S441 cells

S524 was cultured from frozen stock in 5 mL of YS medium until turbid (~72 hr). This culture was transferred to 50 mL fresh YS medium in a long-neck flask and grown for 120 hours. Samples were taken every 24 hours and the cell density (OD_{600nm}) and viral titer (PFU/mL) were measured. The average viral titer at each time point was calculated from plaque assays done in triplicate on cell-free supernatants.

Stably-infected cells were grown from frozen stock and used to purify virus for low MOI experiments (see below). Cultures were incubated for 72 hours and centrifuged at 6,000 x g (Eppendorf 5810R) for 10 min. The virus-containing

supernatant was collected and the titer was measured by plaque assay. The same protocol was used to purify SSV1(*e178*::Topo) from cultures of S538 and SSV1Δ*d244*(*e178*::Topo) from cultures of S530. SSV9-infected cultures were prepared as detailed elsewhere (appendix A) and SSV9 was harvested after 48 hours of incubation at 75º C.

Table 5-1: Strains, viruses, and plasmids used in this work

Strain/vector	Description	Reference
S. solfataricus S441	Novel *Sulfolobus* isolate	Chapter 3
S. solfataricus S524	S441 infected with SSV1 wild-type	This work
S. solfataricus S538	S441 infected with SSV1(e178::Topo)	This work
S. solfataricus S530	S441 infected with SSV1*Δd244*(e178::Topo)	This work
S. solfataricus S506	S441 infected with SSV9 wild-type	This work
SSV1	Wild-type SSV1	Yeats et al., 1982
SSV9	Wild-type SSV9	Wiedenheft et al., 2004
SSV1(*e178*::Topo)	SSV1 shuttle vector with TOPO PCR Blunt II in ORF *e178* (aka pAJC97)	Clore and Stedman 2006
SSV1Δ*d244*(e178::Topo)	SSV1(*e178*::Topo) with ORF *d244* deleted	Iverson and Stedman 2012 (Chapter 2)

Low MOI infections and one-step growth curve experiments

A mid-logarithmic phase culture of S441 was diluted with YS medium to OD_{600nm} of ~0.13 and allowed to recover at 75º C for 2 – 4 hours until the OD_{600nm} reached 0.16, a value experimentally determined to equal 10^8 *Sulfolobus* cells/mL (Drummond 2010; Prangishvili et al., 1999). 30 – 80 mL of S441 cells (~3.0 – 8.0 x 10^9 cells) were added to a corresponding volume of virus calculated to yield the desired MOI. The volume of virus added depended on the titer (5 –

40 mL) and the MOI desired. Negative controls were prepared using an equal volume of YS medium in place of virus. Following addition of virus (or media), cultures were incubated for 2 hours at room temperature with shaking followed by 2 hours at 75º C with shaking. Cultures were then centrifuged at room temperature for 10 min at 3,500 x g (Eppendorf 5810R). The supernatant was discarded and the cells were resuspended in the original culture volume (30 – 80 mL) of YS medium. Cells were centrifuged for another 10 min at 3,500 x g, after which the supernatant was removed and the cells were re-suspended in the original culture volume (30 – 80 mL) of YS medium. Cultures were added to long-neck Erlenmeyer flasks and incubated with shaking at 75º C. The cell density and viral titer were measured at indicated time points throughout the experiment. Viral titers were measured by plaque assay using cell-free supernatants with a minimum of three replicates for each time point. Samples (1 mL) were prepared for plaque assay by pelleting the cells via centrifugation (3 min) at 15,000 x g and carefully collecting the virus supernatant (~500 μL) by pipet. Supernatants could be stored for at least 60 days without a measurable decrease in titer (data not shown; Drummond 2010).

Low MOI infections (0.1) were performed initially to produce high viral titers that were required for higher MOI (5.0) experiments. Low MOI experiments indicated that the peak titer of SSV1 [and SSV1(*e178*::Topo)] occurred between 16 and 42 h.p.i. and purification within this time span routinely yielded titers ~10^9 PFU/mL. The optimal time for SSV1Δ*d244*(*e178*::Topo) purification was 24 – 42

h.p.i. and yielded similarly high titers to SSV1 and SSV1(*e178*::Topo). SSV9 was purified via a different protocol (see appendix A). The virus sample used in each one-step growth curve was used in only a single experiment (i.e. every experiment used a unique viral sample due to the amounts needed) unless otherwise noted. For the growth curves shown in Figure 5-5, the negative control and SSV1Δ*d244*(*e178*::Topo) curves display the average of three independent experiments while the SSV1 and SSV1(*e178*::Topo) curves display the average of two independent experiments.

Calculation of virus burst size

The viral burst size in one-step growth curves was calculated by the following equation:

Equation 1:

$$\frac{\text{maximum PFU } (t_{<24}) - \text{initial PFU } (t_0)}{\text{initial number of cells } (t_0)} = \frac{\text{PFU}}{\text{infected cell}}$$

In the equation above, the maximum PFU ($t_{<24}$) was calculated by identifying the maximum titer within the first 24 h.p.i. and calculating the total number of virus (PFUs). The initial titer is equal to the amount of virus present immediately after culture is placed in 75° C incubator (i.e. the start of the growth curve). The initial number of cells is the number of cells that were initially incubated with the virus for the one-step growth curve (6.0×10^9 cells for all one-step growth curves). As indicated by the above equation, virus burst size is reported in units of PFU/infected cell.

Experiments with "spent" growth medium

Experiments using "spent" YS medium (i.e. media collected from a late-log uninfected S441 culture) in place of fresh YS medium were performed to investigate the effect of spent media on cell growth. Spent media was isolated from uninfected S441 cells in the late logarithmic phase of growth (OD_{600nm} ~ 0.7 – 0.9) via the same protocol used for the separation of virus. Spent media was stored at 4º C until needed (7-10 days). To test the effect of the spent media, 30 mL of S441 cells were prepared as if for a one-step growth curve, however, 18.75 mL of spent media was added instead of virus (see above). A negative control was prepared simultaneously using 18.75 mL of fresh YS in place of spent media. Both cultures were incubated, washed, and grown analogously to one-step growth curve experiments. The cell density (OD_{600nm}) of duplicate cultures was averaged and plotted against time.

UV inactivation of SSV1(e178::Topo) *virions*

SSV1(*e178*::Topo) virions were inactivated with UV light to test the effect of inactivated virus on the growth of S441. 15 mL of titered SSV1 supernatant was added to a polystyrene petri dish (15 x 150 mm) and exposed to 1 J/cm^2 of UV-irradiation using a CL-1000 UV cross-linker (UVP, Inc.) as in Bautista et al. (2015). The titer of UV-irradiated SSV1(*e178*::Topo) was measured following the initial dose and compared to a non-irradiated control. The initial dose reduced SSV1(*e178*::Topo) infectivity ~1000-fold (from 10^8 to 10^5 PFU/mL) so a second

dose (1 J/cm^2) was administered. Again, SSV1(*e178*::Topo) infectivity was reduced by about 1000-fold (to ~10^1 PFU/mL) but not eliminated so a third dose was administered. This third dose appeared to eliminate SSV1(*e178*::Topo) infectivity as the titer was below the detectable limit. To ensure SSV1(*e178*::Topo) was non-infectious, a 4th dose was administered. The total dose of radiation delivered was 4 J/cm^2 (per 15 mL of virus sample).

The effect of UV-inactivated SSV1(*e178*::Topo) virus was tested by substituting the UV-inactivated virus for infectious virus in a one-step growth curve. 18.75 mL of UV-inactivated virus was mixed with 30 mL of S441 cells and the sample was processed as indicated above. An experiment using an identical volume of infectious SSV1(*e178*::Topo) virus from the same un-irradiated stock (18.75 mL) was performed and is shown in Figure 5-4. A negative control using 18 mL of YS in place of virus was also prepared. The cell density (OD_{600nm}) of duplicate cultures was averaged and plotted against time. At the end of the experiment (72 h.p.i.), a sample of the culture "infected" with UV-inactivated SSV1(*e178*::Topo) was checked for the presence of virus by plaque assay. The UV-irradiated sample shown in Figure 5-6 is the average of two independent experiments.

Plaque Assay

Virus titers were measured via plaque assay as in Schleper et al. (1992) with modifications suggested by Albers et al., (2006). Briefly, an S441 culture was diluted with fresh YS medium to OD_{600nm} 0.35 – 0.38 and allowed to recover at 75º C until the cell density reached ~0.40 (2 – 4 hours). Cells were centrifuged at 3,000 x g (Eppendorf 5810R) for 10 min and re-suspended in a volume of YS resulting in ten-fold concentration of the cells. 500 μL of 10X concentrated cells were added to a 5 mL test tube and 100 μL of virus dilution was added. Virus dilutions were prepared by ten-fold serial dilutions of a virus stock in YS medium. 5 mL of soft-layer was added and the mixture was quickly poured and evenly distributed on a warm YS Gelrite® plate (see chapter 2). Plates were allowed to solidify at room temperature for 20 min and then were placed in a sealed plastic box with a moist paper towel and incubated for 48 hours. Plaques began to appear after 24 hours but were most visible after 48 hours. To calculate the viral titer, only dilutions yielding clearly-separated plaques (i.e. non-confluent) were used. If plaques were confluent, they were counted as a single PFU. Plates with 10 – 50 plaques were optimal for quantification as they limited confluence of plaques, but plates with higher PFUs were also counted. All one-step growth curve plaque assays were performed at least in triplicate and averaged to calculate the titer and associated standard deviation at a given time point.

Results

Filtration of virus results in a 100-fold reduction in infectivity

SSV1 was initially purified by a two-step protocol. First, SSV1-infected cultures (S524) were centrifuged (6,000 x g; Eppendorf 5810R) to pellet cells and the virus-containing supernatant was collected. The supernatant was then passed through a 0.45 μm filter (Sartorious Minisart® SFCA) to remove remaining cells. Viral titers purified by this method were routinely low ($10^2 – 10^3$ PFU/mL) but it was unknown if this was due to the purification method or intrinsic to the virus. To investigate if the low viral titers were due to the filtration step, SSV1 was isolated by omitting the filtration step and compared to virus isolated with the filtration step included. Viral titers from samples purified by both protocols (+/- filtration) were measured by plaque assay in quadruplicate and compared (Table 5-2). The filtration step resulted in a roughly 100-fold decrease in the viral titer, indicating that filtration does reduce SSV1 infectivity even though individual virions are ~90 x 30 nm (Stedman et al., 2015) and should easily pass through the filter. This experiment was repeated using SSV1(*e178*::Topo) virus and yielded similar results (Table 5-2).

The observed higher titer could be explained by the presence of contaminating cells that would have otherwise been removed by filtration. If some of those cells are infected, they would be indistinguishable from virus in the plaque assay, resulting in an artificially inflated titer. To check this, unfiltered supernatants were incubated for 10 days at 75º C but no growth was observed,

indicating that any potentially contaminating cells are present in low abundance. Furthermore, spotting supernatants on YS plates and incubating for 10 days at 75° C also did not result in visible cell growth. *Sulfolobus* cell-like particles have been observed when analyzing unfiltered virus supernatants in the TEM. However, they are present in low abundance and may not be viable (data not shown). These results suggest that separation of SSV solely by centrifugation is suitable and results in a minimal loss of infectious virus. This protocol was used to isolate all virus used in the experiments below.

Table 5-2: Effect of filtration (0.45 μm) on SSV1 and SSV1(e178::Topo) supernatants

SSV1	**+ Filtration (PFU/mL)**	**No filtration (PFU/mL)**
Plate 1	2.1×10^3	4.1×10^5
Plate 2	5.8×10^3	1.2×10^5
Plate 3	4.3×10^3	2.5×10^5
Plate 4	1.9×10^3	2.7×10^5
Average ± S.D.	**$3.53 \pm 1.86 \times 10^3$**	**$2.63 \pm 1.19 \times 10^5$**

SSV1(e178::Topo)	**+ Filtration (PFU/mL)**	**No filtration (PFU/mL)**
Plate 1	5.1×10^4	1.1×10^6
Plate 2	3.9×10^4	9.0×10^5
Plate 3	1.3×10^4	1.9×10^6
Plate 4	5.5×10^4	2.1×10^6
Average ± S.D.	**$3.95 \pm 1.89 \times 10^4$**	**$1.5 \pm 5.89 \times 10^6$**

Newly transfected Sulfolobus *yields low viral titers*

Because one-step growth curves at a high multiplicity of infection (MOI) require large amounts of virus (~10^9 PFU/mL), it was essential to develop a robust and reliable method for the production of high viral titers. UV-induction of SSV1-infected cultures has been reported to yield titers of up to 10^9 PFU/ml, although induction has been notoriously difficult to reproduce, volumes are

generally low, UV-irradiation seems to vary depending on the host strain used, and finally, UV-irradiation could also induce mutations (Schleper et al., 1992; Fusco et al., 2014; Stedman personal communication). Therefore, UV-irradiation independent means of virus production were investigated. Mitomycin C has been reported as a method to induce SSV1 production, but very large amounts are required, probably due to the instability of mitomycin C at low pH (Cannio et al., 1998; Liu and Li, 2002; Drummond 2010). Initially, *S. solfataricus* strain S441 was transfected with SSV1 by electroporation (See chapter 3). Following electroporation, cells were diluted into 50 mL of YS medium and incubated at 75° C. Samples (1 mL) were removed every 24 hours and the viral titer in cell-free supernatants was measured by plaque assays in triplicate. The SSV1 titer peaked ~72 – 96 hours post-electroporation and did not exceed 10^5 PFU/mL (data not shown). Similar results were observed following electroporation of S441 with the SSV1 shuttle vector SSV1(*e178*::Topo). Glycerol stocks from SSV1- and SSV1(*e178*::Topo)-infected cultures were prepared and used for preparation of virus below.

Stably infected cultures produce reasonably high (10^7 – 10^8 PFU/mL) titers

SSV1-infected cells (S524) were prepared from frozen stock and grown in 5 mL of YS medium until turbid (~72 hours). This culture was transferred to 75 mL and incubated at 75° C for 120 hours. Cell density and viral titer were measured every 24 hours via plaque assay to determine the optimal time for

collection and isolation of virus (Figure 5-1). The initial titer in cell-free supernatants (immediately after dilution) was low (~10^1 PFU/mL) and reached a maximum of 10^7 PFU/mL by 72 hours. The titer then gradually decreased as the cells reached stationary phase. Repetitions of this experiment routinely resulted in viral titers in this range (10^5 – 10^8 PFU/mL). Similar results were obtained for SSV1(*e178*::Topo)- and SSV1Δ*d244*(*e178*::Topo)-infected cultures after 72 hour incubations (data not shown). Although greater viral titers were obtained by this method compared to post-electroporation samples (above), these titers were still not sufficient to perform high MOI one-step growth curve experiments.

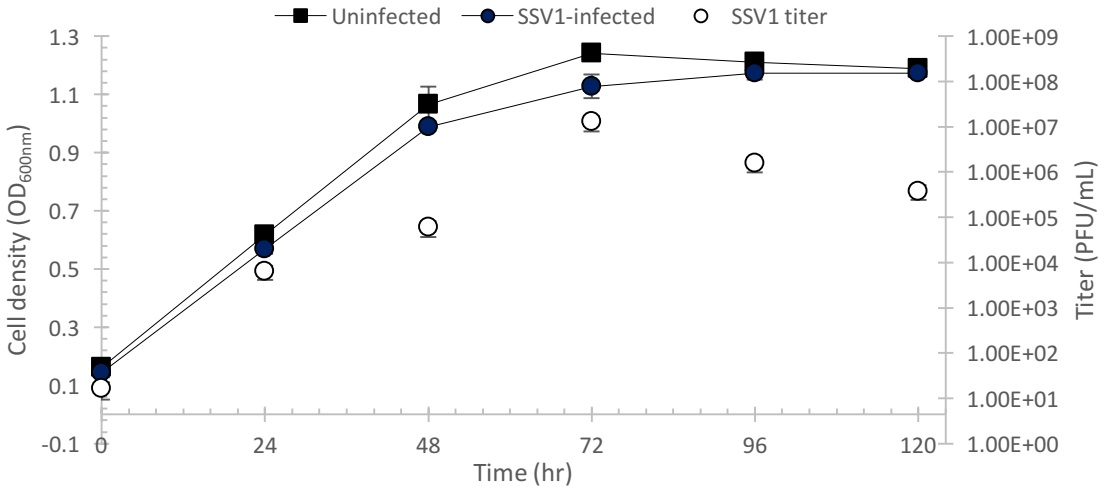


Figure 5-1: Growth and virus production of SSV1-infected S441 cells
Sulfolobus cells of strain S441 stably infected with SSV1 (circles) and uninfected S441 (squares) were diluted to 75 mL with fresh YS medium and incubated for 120 hours at 75° C. Samples (1 mL) were taken every 24 hr to determine cell density (OD_{600nm}) and to measure the viral titer by plaque assay (open circles). The average OD_{600nm} and viral titer (PFU/mL) of two independent cultures are presented with standard deviations. The viral titer was measured in triplicate at each time point, in each replicate culture.

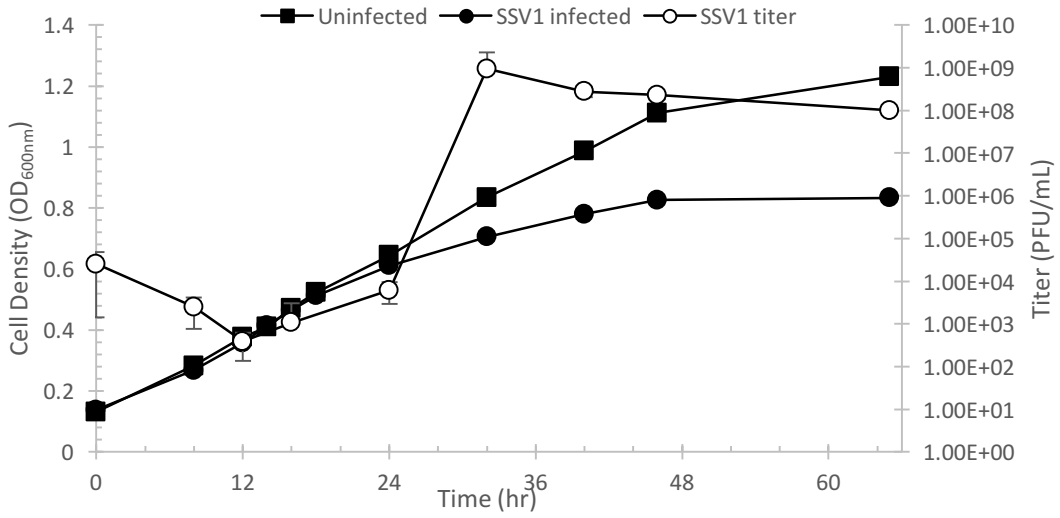
Infections with SSV1 and SSV1(e178::Topo) at low MOI

Infection of S441 cells with SSV1 at a low MOI was initially performed to measure the kinetics of viral infection. SSV1 with a titer of 2.0 x 10^8 PFU/mL was isolated from a 72-hour old culture of S524 (as above) and 4 mL of this virus preparation was mixed with 80 mL of S441 cells (OD_{600nm} = 0.16) for an MOI of 0.1 (Figure 5-2A). Following a 2-hour incubation at room temperature and 2-hour incubation at 75º C, the un-adsorbed virus was removed by centrifugation and the cells were resuspended in fresh YS medium. An uninfected control was simultaneously prepared using an equal volume of YS in place of virus supernatant. Cultures were grown at 75º C with shaking in long-neck flasks. Samples (1 mL) were taken at indicated times to measure cell density (OD_{600nm}) and the SSV1 titer (Figure 5-2A). The cell densities of both the infected and uninfected cultures were nearly identical for the first 16 hours, after which the infected cells began to grow at a decreased rate. The titer of SSV1 decreased by almost two orders of magnitude over the first 12 h.p.i. Between 12 and 16 h.p.i., the SSV1 titer began to increase. A large burst in virus production was observed between 24 and 32 h.p.i., with the titer reaching its maximum of 9.38 x 10^8 PFU/mL. The reduction in growth rate of the infected culture corresponds with this burst in viral titer. The SSV1 titer gradually decreased over the final 48 hours of the experiment, corresponding to the static growth of the infected culture during this time.

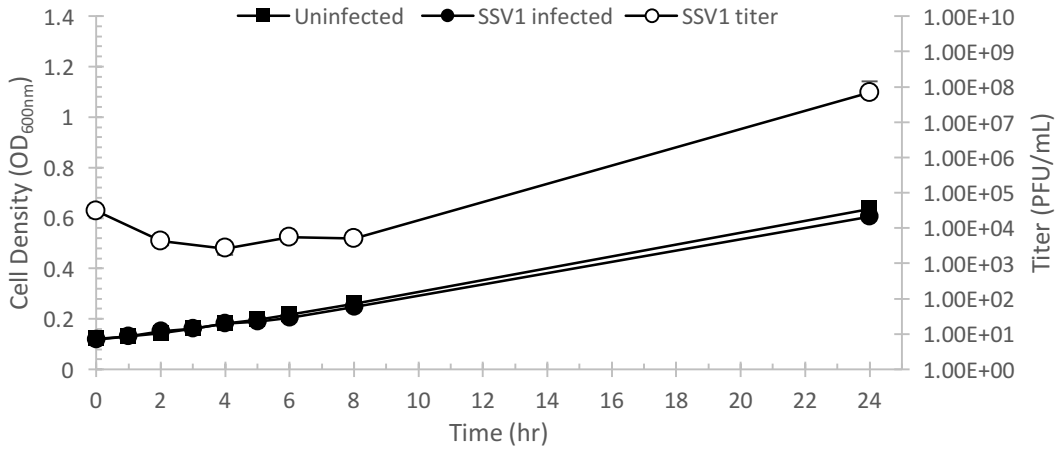
A second infection with SSV1 at the same MOI (0.1) was conducted to more closely measure the SSV1 titer during the first 8 hours of infection (Figure 5-2B). The SSV1 sample used was the same used in the previous experiment (Figure 5-2A). The titer of SSV1 decreased ten-fold within 2 h.p.i. and hovered at this value until increasing almost 10,000 fold between 8 and 24 h.p.i. This increase in titer occurred slightly earlier than the burst observed in the initial experiment (Figure 5-2A). The infected and uninfected cells exhibit almost identical growth curves to those in the initial experiment (Figure 5-2A).

A third experiment using virus produced by the shuttle vector SSV1(*e178*::Topo) at a MOI of 0.1 was performed and displays a similar growth curve to wild type SSV1 with a few notable differences (Figure 5-2C). SSV1(*e178*::Topo) virus with a titer of 1.6×10^8 PFU/mL was isolated from a 72-hour old culture of S538. Five mL of SSV1(*e178*::Topo) virus was mixed with 80 mL of S441 culture ($OD_{600nm} = 0.16$). Unlike the SSV1 growth curves, the titer of SSV1(*e178*::Topo) did not decrease during the first 12 h.p.i. but instead increased almost 100-fold. After this initial increase, a slight decrease in titer was observed until 24 h.p.i followed by a 100,000-fold increase that reached a maximum of 3.0×10^9 PFU/mL by 40 h.p.i. Overall, data from the growth curves indicates that infection with SSV1, or SSV1(*e178*::Topo), at a low MOI results in a burst of virus production between 16 and 32 h.p.i. that can reach titers $>10^9$ PFU/mL. This trend was exploited to produce the greater titers needed for higher MOI experiments.

A


Uninfected
SSV1 infected
SSV1 titer
Cell Density (OD_{600nm})
Titer (PFU/mL)
Time (hr)
0
0.2
0.4
0.6
0.8
1
1.2
1.4
0
12
24
36
48
60
1.00E+00
1.00E+01
1.00E+02
1.00E+03
1.00E+04
1.00E+05
1.00E+06
1.00E+07
1.00E+08
1.00E+09
1.00E+10

B


Uninfected
SSV1 infected
SSV1 titer
Cell Density (OD_{600nm})
Titer (PFU/mL)
Time (hr)
0
0.2
0.4
0.6
0.8
1
1.2
1.4
0
2
4
6
8
10
12
14
16
18
20
22
24
1.00E+00
1.00E+01
1.00E+02
1.00E+03
1.00E+04
1.00E+05
1.00E+06
1.00E+07
1.00E+08
1.00E+09
1.00E+10

C

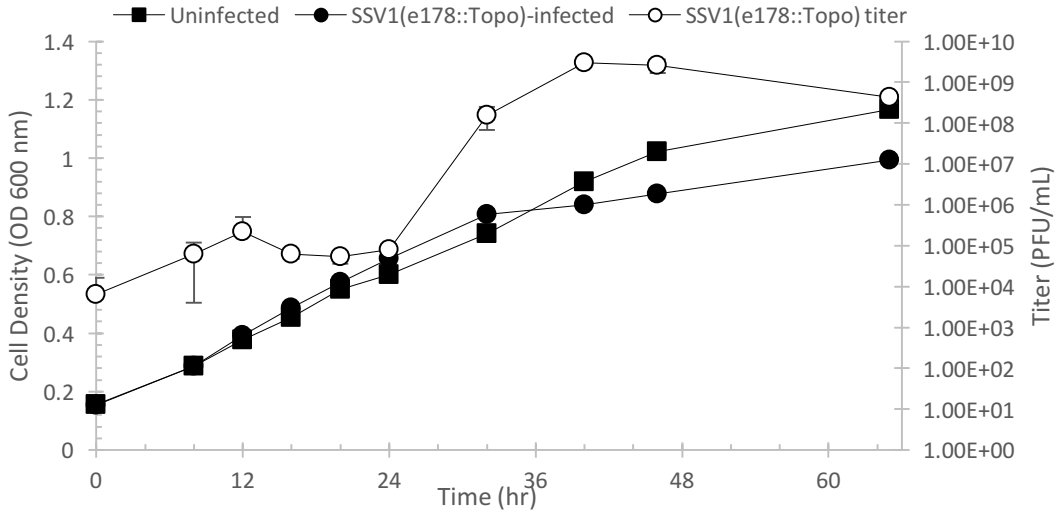

Uninfected
SSV1(e178::Topo)-infected
SSV1(e178::Topo) titer
Cell Density (OD 600 nm)
Titer (PFU/mL)
Time (hr)
0
0.2
0.4
0.6
0.8
1
1.2
1.4
0
12
24
36
48
60
1.00E+00
1.00E+01
1.00E+02
1.00E+03
1.00E+04
1.00E+05
1.00E+06
1.00E+07
1.00E+08
1.00E+09
1.00E+10

Figure 5-2: Growth curves of S441 cells infected with SSV1 and SSV1(*e178*::Topo) at low MOI
(A) Growth curve of S441 cells infected with SSV1 wild-type virus at a MOI of 0.1. 80 mL of S441 culture (8.0×10^9 cells) were mixed with 4 mL of SSV1 with a titer of 2.0×10^8 PFU/mL (8.0×10^8 PFU) as detailed in methods (closed circles). A negative control was simultaneously prepared using 80 mL of S441 culture and 4 mL of YS medium in place of virus (closed squares). The cell density (OD_{600nm}) and SSV1 titer (PFU/mL) were measured at indicated times in the figure. The SSV1 titer (open circles) at each time point represents the average titer calculated by plaque assay in triplicate (standard deviation indicated by bars).
(B) Growth curve of S441 cells infected with SSV1 wild-type virus at a MOI of 0.1. The experiment in panel A was repeated with a focus on the initial 24 hours of infection. Both infected (closed circles) and uninfected cultures (closed squares) were prepared in an identical fashion to those in panel A and used the same SSV1 sample. The cell density (OD_{600nm}) and SSV1 titer (PFU/mL) were measured at indicated times in the figure. The SSV1 titer (open circles) at each time point represents the average titer calculated by plaque assay in triplicate (standard deviation indicated by bars).
(C) One-step growth curve of S441 cells infected with SSV1(*e178*::Topo) shuttle vector at a MOI of 0.1. 80 mL of S441 culture (8.0×10^9 cells) were mixed with 5 mL of SSV1(*e178*::Topo) with a titer of 1.6×10^8 PFU/mL (8.0×10^8 PFU) and prepared identically to other samples (closed circles). A negative control was simultaneously prepared with 80 mL of S441 culture and 5 mL of of YS medium added in place of virus (closed circles). The cell density (OD_{600nm}) and SSV1(*e178*::Topo) titer (PFU/mL) were measured at the indicated times. The SSV1(*e178*::Topo) titer (open circles) at each time point represents the average titer calculated by plaque assay in triplicate (standard deviation indicated by bars).

Infection with SSV9 at low MOI

An infection of *Sulfolobus* with SSV9 at an MOI of 0.1 was also performed and displays some notable differences with the SSV1 low MOI experiments. SSV9 with a titer of 1.3×10^8 PFU/mL was isolated from infected S441 cells prepared as detailed elsewhere (see appendix A). 8 mL of SSV9 supernatant was added to 80 mL of S441 cells ($OD_{600nm} = 0.16$) and prepared as done previously for SSV1 low MOI infections. The cell density was measured at indicated times and the average SSV9 titer was determined from cell-free supernatants by plaque assays in triplicate (Figure 5-3). Within 8 h.p.i., the titer

of SSV9 increased almost 100-fold and quickly reached a maximum of 3.6 x 10^7 PFU/mL by 16 h.p.i. The titer then gradually decreased until 40 h.p.i. and appears to level off until the end of the experiment (65 h.p.i.), finishing at ~10^6 PFU/mL. One striking difference between the SSV9 and SSV1 infections at 0.1 MOI was the growth of the infected cells. The cell density of the SSV9-infected culture mirrors that of the uninfected culture during the first 12 h.p.i. but quickly stagnates, barely increasing until approximately 40 h.p.i. This growth inhibition emulates what is observed in higher MOI infections with SSV1 (Figure 5-5). Additionally, cell debris was observed in the culture after 48 h.p.i.

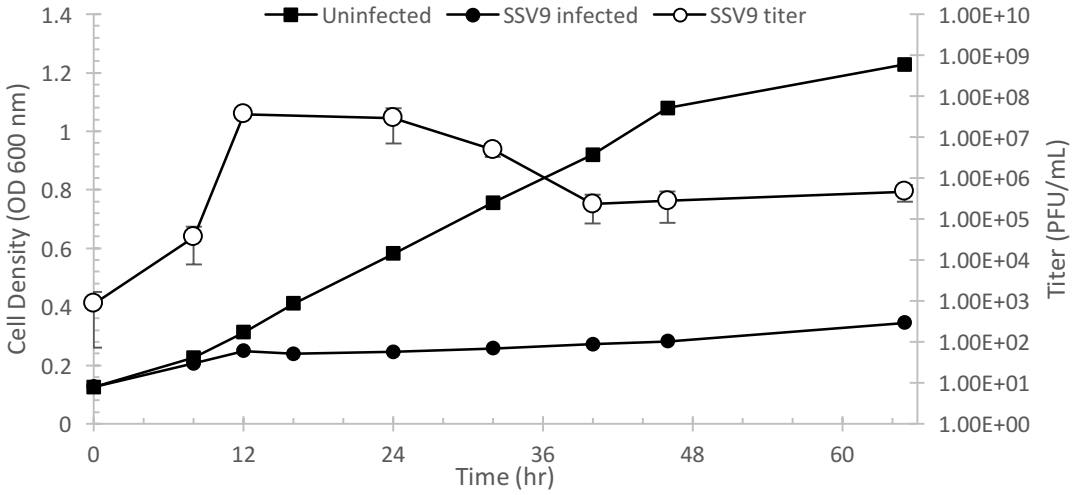


Figure 5-3: Growth curve of S441 infected with SSV9 at a MOI of 0.1
80 mL of S441 cells (8.0 x 10^9 cells) were mixed with 8 mL of SSV9 with a titer of 1.3 x 10^8 PFU/mL (1.04 x 10^9 PFU) and prepared as described in methods (closed squares). A negative control was likewise prepared with 80 mL of S441 cells and 8 mL of YS added in place of virus (closed circles). The cell density (OD_{600nm}) and SSV9 titer (PFU/mL) were measured at the indicated times. Plaque assays in triplicate were performed at each time point to calculate the average SSV9 titer (standard deviation indicated by bars).

One-step growth curves at a MOI of 5

The observation that high SSV1 and SSV1(*e178*::Topo) titers (~10^9 PFU/mL) were produced by infection at a low MOI enabled experiments with a higher MOI. An infection of S441 with SSV1(*e178*::Topo) at an MOI of 0.1 produced a sample of SSV1(*e178*::Topo) virus with a titer of 8.0 x 10^8 PFU/mL after 40 h.p.i. 18.75 mL of this sample (1.5 x 10^{10} PFU) was used to infect 3.0 x 10^9 cells of S441 (30 mL) at an MOI of 5.0 (Figure 5-4). Unlike in previous low MOI experiments (Figure 5-2), cells infected at the higher MOI exhibit a decreased rate of growth relative to the negative control and are reminiscent of cells infected with SSV9 at a low MOI (Figure 5-3). The titer of SSV1(*e178*::Topo) displays a similar pattern to the low MOI SSV1(*e178*::Topo) infection (Figure 5-2C). The SSV1(*e178*::Topo) titer increases almost three orders of magnitude within the first 12 hours (to ~10^7 PFU/mL), similar to the low MOI infection with SSV1(*e178*::Topo) but not SSV1 (Figure 5-2). Following the initial increase, the titer of SSV1(*e178*::Topo) remains relatively constant at ~10^7 PFU/mL before reaching a maximum of 10^8 PFU/mL (24 h.p.i.). The titer remained fairly constant (~10^7 PFU/mL) for the remainder of the experiment (42 h.p.i.).

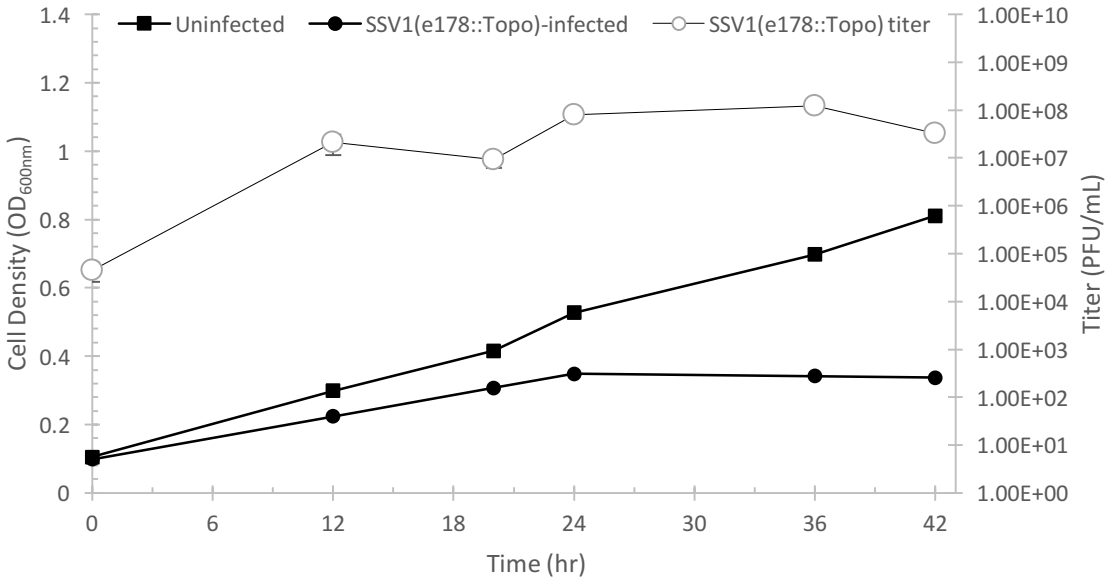


Figure 5-4: One-step growth curve of S441 infected with SSV1(*e178*::Topo) at a MOI of 5

30 mL of S441 cells (3.0 x 10^9 cells) were mixed with 18.75 mL of SSV1(*e178*::Topo) at a titer of 8.0 x 10^8 PFU/mL (1.5 x 10^{10} PFU) and prepared as described for other growth curves (closed circles). A negative control was prepared simultaneously by adding 18.75 mL of YS medium in place of virus to 30 mL of cells (closed squares). The cell density (OD_{600nm}) and SSV1(*e178*::Topo) titer (PFU/mL) were measured at indicated times. Plaque assays in triplicate were performed on cell-free supernatants at each time point to calculate the average viral titer (standard deviation indicated by bars).

One-step growth curves of SSV1, SSV1(e178::Topo) and the SSV1 mutant SSV1Δd244(e178::Topo)

One-step growth curves were performed to compare the infection kinetics of SSV1(*e178*::Topo), SSV1Δ*d244*(*e178*::Topo) and wild-type SSV1 (Table 5-1). High-titer samples of each virus were isolated by first infecting a culture of S441 at low MOI (~0.1) and purifying the virus between 32 and 40 h.p.i., analogous to what was done previously for SSV1(*e178*::Topo) (Figure 5-2C, see methods). Purified virus was used to infect S441 at a MOI of 5 as was done in previously (Figure 5-4). The negative control and SSV1*Δd244*(*e178*::Topo) growth curves were performed in triplicate and SSV1 and SSV1(*e178*::Topo) growth curves

were performed in duplicate. Cell density (OD_{600nm}) and viral titer were measured at the indicated times. The average viral titer at each time point was determined by plaque assays in triplicate (Figure 5-5). SSV1- and SSV1(*e178*::Topo)-infected cells exhibited similar patterns of growth (Figure 5-5A). Interestingly, SSV1Δ*d244*(*e178*::Topo)-infected cells display a noticeably different growth phenotype from cells infected with either of the other viruses (Figure 5-5B). SSV1Δ*d244*(*e178*::Topo)-infected cells appear indistinguishable from wild type at 12 h.p.i., but then appear to grow at an accelerated rate relative to the uninfected control. The rate of cell growth begins to decrease between 24 and 36 h.p.i. and the cell density decreases between 36 and 44 h.p.i. before leveling out. While cell debris was observed after 65 h.p.i. in all infected cultures, debris was visible in SSV1Δ*d244*(*e178*::Topo)-infected cultures within 36 h.p.i. Samples of SSV1Δ*d244*(*e178*::Topo)-infected cells between 12 and 36 h.p.i. were observed under the light microscope to check for any abnormally large cells, which may explain the increase in cell density that was observed during these times relative to the uninfected control. However, the cells appeared indistinguishable from the uninfected controls.

The viral titer increased within the first 12 h.p.i. for all viruses tested (Figure 5-5). Both the SSV1 and SSV1(*e178*::Topo) titer increased almost 1000-fold whereas SSV1Δ*d244*(*e178*::Topo) increased by approximately 100-fold. After reaching a maximum titer of 4.9×10^8 PFU/mL after 16 h.p.i., the SSV1(*e178*::Topo) titer steadily decreases throughout the course of the

experiment, similar to low MOI growth curves. The titer of SSV1 somewhat mirrors that of SSV1(*e178*::Topo), although a ten-fold increase in titer is observed at ~36 h.p.i., after which the titer remains relatively constant for the duration of the experiment. After an initial increase, the SSV1Δ*d244*(e178::Topo) titer dips roughly 10-fold between 16 and 24 h.p.i. The titer then undergoes a second increase (~100-fold) between 24 and 40 h.p.i., reaching a maximum of 1.84 x 10^9 PFU/mL. This second burst correlates with the observed reduction in the growth rate of the host cells.

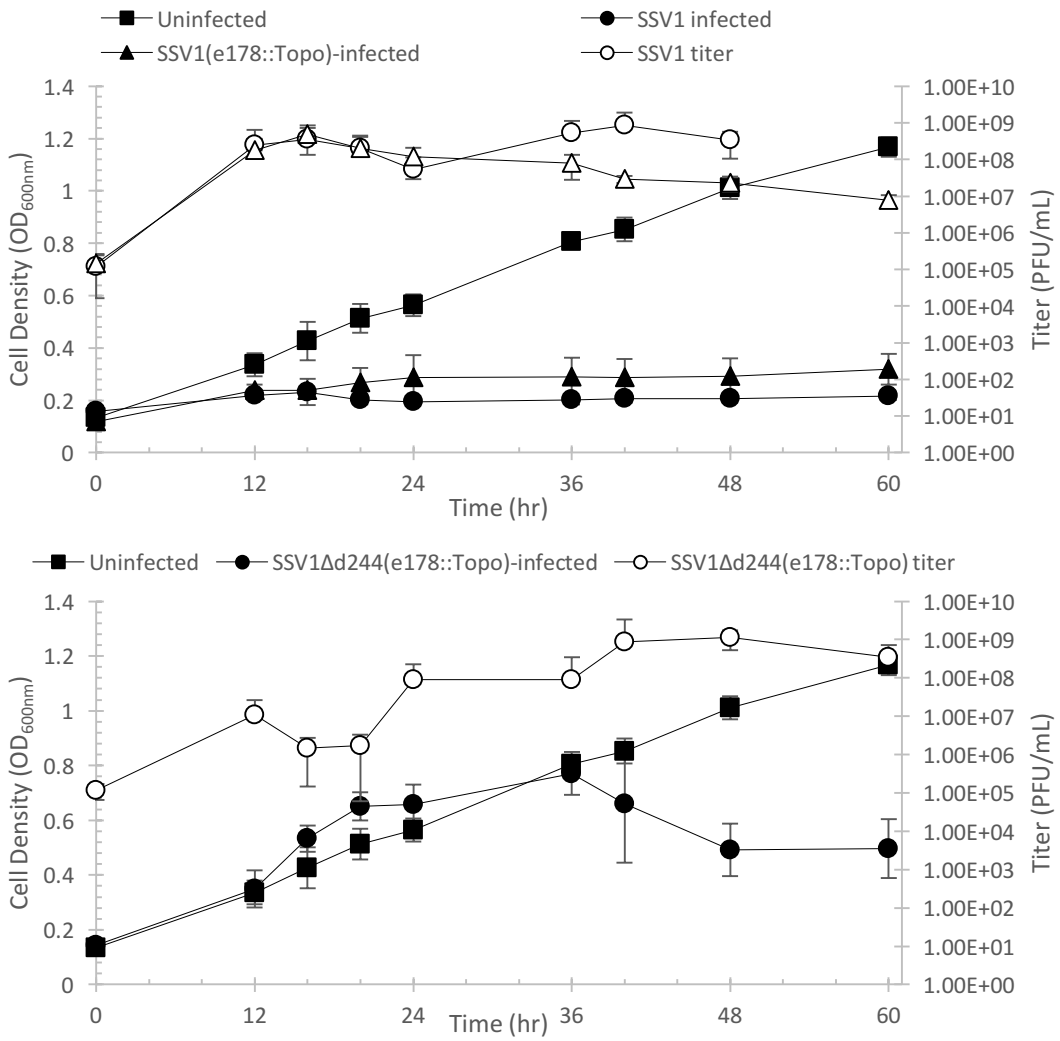


Figure 5-5: One-step growth curve of S441 cells infected with SSV1, SSV1(*e178*::Topo), and SSV1Δ*d244*(*e178*::Topo) at a MOI of 5
60 mL of S441 culture (6.0 x 10^8 cells) were mixed with SSV1 (panel A), SSV1(*e178*::Topo) (panel A), and SSV1Δ*d244*(*e178*::Topo) (panel B) virus at a MOI of 5. The amount of virus added to each sample varied depending on the initial titer of the virus, but the final volume was adjusted to 40 mL with YS medium. A negative control was prepared by adding 40 mL of YS medium in place of virus. The same data for the negative control is displayed in panels A and B. The cell density (OD_{600nm}) and viral titer (PFU/mL) of each sample were measured at indicated times. Negative control and SSV1Δ*d244*(*e178*::Topo) growth curves were performed in triplicate while SSV1 and SSV1(*e178*::Topo) growth curves were performed in duplicate. The average cell density from replicate experiments is displayed with associated standard deviation. The average viral titer at each time point is displayed with standard deviations and was calculated from triplicate plaque assays performed for each experimental replicate.

The initial 24 hours of the one-step growth curve were used to calculate the burst size for each of the viruses analyzed. The burst size of a virus is defined as the number of virus produced per infected cell and was calculated by equation 1 (see methods). The burst size of each virus is presented in the table 5-3.

Table 5-3: Virus burst sizes

Virus	Maximum PFU (t<24)	Initial PFU (t0)	Burst size (PFU/cell)
SSV1	2.11 x 10^{10}	7.20 x 10^6	3.52
SSV1(e178::Topo)	2.95 x 10^{10}	8.52 x 10^6	4.92
SSV1Δ*d244*(e178::Topo)	6.60 x 10^8	7.02 x 10^6	0.11

The effects of spent media and UV-inactivated virus on S441 growth

Due to the method of virus separation used here, it was important to investigate the effect of used (aka "spent") *Sulfolobus* media in the one-step growth curves to ensure that the effects observed in the growth curves were due to the virus and not some other factor present in the medium. We were also interested in observing the effect of UV-inactivated viral particles

[SSV1(*e178*::Topo)] on cell growth as it has been reported that UV-inactivated SSV9 at a low MOI causes population-wide stasis in *S. islandicus* (Bautista et al., 2015). An experiment similar to the one-step growth curves presented in figures 5-4 and 5-5 was performed by adding 18.75 mL of UV-inactivated SSV1(*e178*::Topo) or medium from a stationary phase S441 culture in place of infectious virus to 30 mL of S441 cells (~3.0 x 10^8 cells). The amount of UV-inactivated virus added was equal to an MOI of 5 prior to UV-inactivation. Spent S441 medium was harvested from a culture in the late logarithmic growth phase via centrifugation, as used to isolate virus. SSV1(*e178*::Topo) was UV-inactivated by irradiating a 15 mL sample in a polystyrene petri dish (15 x 150 mm) with four doses of 1 J/cm^2 using a UV cross-linker (UVP). Plaque assays on the UV-inactivated virus showed that each dose lowered the infectivity approximately 1000-fold and that infectivity was below the detectable limit after three doses of irradiation. A fourth dose was administered to ensure inactivation of virus was complete. The UV-inactivated SSV1(*e178*::Topo) virus was from the same stock as the infectious SSV1(*e178*::Topo) virus used in a previous experiment (Figure 5-4). A negative control was prepared using 18.75 mL of YS medium in place of virus or spent medium.

Following the addition of virus or media, cells were prepared in an analogous fashion to one-step growth curve experiments. Samples (1 mL) were taken at indicated times and the cell density (OD_{600nm}) from duplicate experiments was plotted (Figure 5-6). Cells that were incubated with UV-

inactivated SSV1(*e178*::Topo) grew at a rate that was indistinguishable from uninfected cells. Cultures receiving spent media grew similarly. These data suggest that neither UV-inactivated virus nor the spent media has a significant effect on cell growth. The culture receiving UV-inactivated virus was confirmed to be virus-free by plaque assay at the conclusion of the experiment.

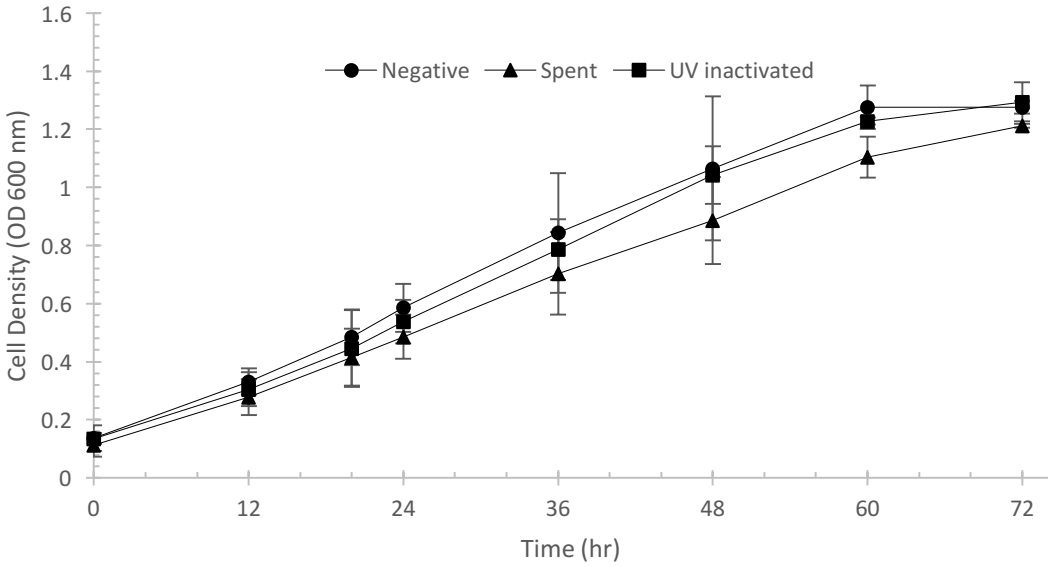


Figure 5-6: The effect of spent growth medium and UV-inactivated SSV1(*e178*::Topo) on S441 cell growth
3.0×10^8 S441 cells (30mL) were mixed with 18.75 mL of UV-inactivated SSV1(*e178*::Topo) or spent YS medium to test the effect of each on cell growth. Spent YS medium was collected from uninfected S441 cells grown to late log phase (OD_{600nm} ~0.8) and isolated by centrifugation (see methods). SSV1(*e178*::Topo) was UV-inactivated and confirmed to be non-infectious via plaque assay prior to the experiment. A negative control using 18.75 mL of YS medium in place virus/spent media was also prepared. All cultures were prepared identically to one-step growth curves (Figures 5-4 and 5-5). The average cell density (OD_{600nm}) of duplicate growth experiments is displayed along with standard deviations. The culture receiving UV-inactivated virus was shown to be virus-free by plaque assay at 72 h.p.i., as expected. UV-inactivated SSV1(*e178*::Topo) was from the same stock used in the one-step growth curve in Figure 5-4.

Discussion

Filtration of SSV1 results in a decrease in infectivity

Initially, purification of SSV1 was performed by centrifugation of infected cultures followed by filtration of the supernatant through a 0.45 μm syringe filter. Omission of the filtration step resulted in an almost 100-fold increase in the viral titer, illustrating that filtration reduces SSV1 infectivity (Table 5-2). Similar results were observed for SSV1(*e178*::Topo) virus. Analysis of virus supernatants isolated without filtration by TEM showed them to be largely devoid of contaminating cells, suggesting that isolation of SSV1 by centrifugation is suitable. Why filtration reduces infectivity is unclear. It is possible that the viral particles are damaged when passed through the filter and lose infectivity. Alternatively, viral particles may become bound to the filter membrane and be removed from the solution. It is also conceivable that many virus particles are bound to each other, causing the aggregate to be larger than the pore size of the membrane (Quemin et al., 2015; Stedman et al., 2003; Redder et al., 2009).

SSV1 virions, like most fuselloviruses, commonly form rosettes due to interactions between the tail fibers of different virions (Redder et al., 2009; Wiedenheft et al., 2004). SSV1 is also commonly observed bound to cell debris and vesicles (Martin et al., 1984; personal observations). The infectivity of these SSV1 aggregates is unknown, but the prevalence of these aggregates suggests that viral titers likely underestimate the number of infectious virions in a given

solution. It was recently shown that treatment with NaCl (1.5 M) and ethanol (1%) results in partial dissociation of large SSV1 aggregates by TEM analysis, although ethanol treatment reduced virus infectivity (Quemin et al., 2015). Similar treatments may yield increased viral titers, although the utility of this in the purification of SSV1 remains to be investigated.

Newly transfected Sulfolobus *results in the production of low viral titers*

The viral production in cultures transfected with SSV1 DNA was measured and found to be relatively low, only reaching maximum titers of ~10^5 PFU/mL by 72 hours post transfection. The transformation efficiency of *Sulfolobus* via electroporation is known to be quite low (10^{-5} transformants/μg) and probably explains the low titers observed following transfection (Schleper et al., 1992). Because only a very small percentage of the population becomes infected with viral DNA, this limits the total number of virus-producing cells. Following electroporation, the culture gradually generates more virus which are then available to infect the rest of the uninfected population. Although the titer appears low, the number of infected cells is probably quite high and may explain why revitalization of a previously transfected culture (that previously produced low titers) results in increased virus production (see below and Figure 5-1).

Stably infected cultures produce moderately high (10^7 – 10^8 PFU/mL) titers

SSV1 cultures that had been transfected via electroporation were frozen and later used as a source for the production of virus. When these cultures were

re-grown from frozen stock, SSV1 titers often reached values of 10^7 –10^8 PFU/mL, significantly higher than titers from freshly transfected cells but still not high enough to prepare virus for one-step growth curve experiments (Figure 5-1). Similar results were obtained for the SSV1 shuttle vector, SSV1(*e178*::Topo). Although SSV1-infected cells grew at a slightly depressed rate relative to uninfected controls, these cultures still grew to high cell densities. It is not unusual for SSV1-infected *Sulfolobus* cultures to exhibit growth rates similar to that of an uninfected culture (Fusco et al., 2013; Contursi et al., 2006).

Infection of Sulfolobus *at a low MOI*

Sulfolobus S441 was initially infected at a low MOI because the high titers required for one-step growth curves could not be obtained. Fortuitously, low MOI infections resulted in a burst of virus production that reached titers in excess of 10^9 PFU/mL (Figure 5-2A). Similar results were obtained for low MOI infections of SSV1(*e178*::Topo) (Figure 5-2C) and SSV1Δ*d244*(e178::Topo) (Data not shown). Cells infected at low MOI displayed growth rates similar to the uninfected controls for the first 16 – 24 h.p.i., after which a noticeable growth defect was observed that correlated with a dramatic increase in the viral titer.

High titers of SSV1 (>10^8 PFU/mL) have been difficult to isolate. The highest reported titers (~10^9 PFU/mL) for SSV1 have all required irradiation with UV-light to induce viral replication in the host (Schleper et al., 1992; Fusco et al., 2015a). A self-stated standardized protocol for UV-induction has been

published, however, the reliability of this is unknown (Fusco et al., 2015a). Our results indicate that infection of strain S441 with SSV1 at a low MOI results in viral titers >10^9 PFU/mL within 40 hours of infection, equaling or exceeding the titers reported for UV-induction in other hosts (Figure 5-2). This protocol was also successfully applied to the SSV1 shuttle vector SSV1(*e178*::Topo) and an SSV1 mutant lacking the *d244* gene, showing it is reliable and robust. It is unclear how successful this protocol is with other fuselloviruses, as similarly high titers (~10^9 PFU/mL) of SSV9 could not be isolated in a similar fashion (Figure 5-3).

The steady titer exhibited by SSV1 during the initial 12 h.p.i. correlates well with a typical latent period of a viral infection (Figure 5-2A; Ellis et al., 1939; Drummond 2010). What is unusual is the titer of SSV1(*e178*::Topo) exhibits a ~100-fold increase in titer during this same time frame. The only known difference between the two viruses is that the genome of SSV1(*e178*::Topo) harbors a bacterial plasmid inserted into ORF *e178* (Clore and Stedman 2006). It is probable that the data reflect a limited sample size and could represent an experimental artifact. In support of this, one-step growth curves with SSV1 do not display this trend and are nearly identical to those of SSV1(*e178*::Topo) performed at the same MOI (Figure 5-5A).

One-step growth curves in Sulfolobus *at a MOI of 5*

SSV1, SSV1(*e178*::Topo), and SSV1Δ*d244*(*e178*::Topo) viruses were isolated from low MOI infections and subsequently used in one-step growth curves at a MOI of 5. MOIs in the range of 5 – 10 theoretically result in >99% of the cells encountering at least one infectious viral particle, suggesting that a synchronous infection was achieved in these experiments.

Comparisons of one-step growth curves

Each infection at high MOI resulted in a large burst of virus production within 12 h.p.i. that was not observed in the low MOI infections (Figure 5-5). This difference is likely due to a greater initial number of infected cells at the higher MOI, resulting in increased virus production over a shorter time span. The overall pattern of virus production in high MOI infections with SSV1 and SSV1(*e178*::Topo) was reminiscent of the low MOI infections (Figure 5-5A). Following the initial burst in virus production, the titers in both samples remain relatively constant throughout the course of the experiment, although SSV1 achieves a greater maximum titer. Preliminary data from our own lab, as well as published data elsewhere, indicate that SSV1 loses infectivity when exposed to high temperature and low pH (Schleper et al., 1992; Drummond 2010). This instability makes it unlikely that the steady-state titer is due to virus that persists in the medium for an extended period of time (i.e. following the initial burst, virus production ceases). Rather, it is likely the result of continuous production by

infected cells. The pattern of virus production in SSV1Δ*d244*(e178::Topo) cultures was slightly different (Figure 5-5B). Following an initial burst, the titer of SSV1Δ*d244*(e178::Topo) remains roughly constant until an increase in virus production occurs at ~20 h.p.i. that yields a maximum titer of ~10^9 PFU/mL at 40 h.p.i. This increase in virus production corresponds with a reduction in the cell density and the appearance of cell debris in the culture. Virus production in SSV1Δ*d244*(*e178*::Topo) cells occurs more gradually compared to SSV1 and SSV1(*e178*::Topo), although the maximum titers achieved by all three infected cultures are similar (~10^9 PFU/mL).

Virus burst size and latent period

The data provided by a one-step growth curve enables the determination of several infection parameters including viral burst size and virus latent period. The burst size is defined as the number of virus produced per infected cell and was calculated by equation 1 (see methods). The maximum titer within the first 24 h.p.i. was used to calculate the burst size as this was likely to reflect the amount of virus produced following a single cycle of virus replication. In support of this time scale, it was shown that the transcription cycle of SSV1 following UV-induction requires ~8.5 hours for completion (Fröls et al., 2007). Although the transcription cycle following synchronous infection is not known, it is likely longer. The initial titer (0 hour) was then subtracted from the maximum titer and this

value was divided by the total number of cells that were initially incubated with virus (6.0×10^9 cells).

Our data indicate low burst sizes for each of the three viruses analyzed (See equation 1 and Table 5-3). The low values may indicate that a synchronous infection was not achieved, as a burst size of 1 PFU/cell is equal to one virus produced per infected cell. Because we cannot be sure that a synchronous infection was achieved, it is not possible to make an accurate determination of the true SSV1 burst size at this time. However, the data do allow for comparisons to be made between the viruses used in this work, though, as they were all treated under the same experimental conditions. The amount of virus production of SSV1 and SSV1(*e178*::Topo) appear to be very similar, as is expected. The amount of SSV1Δ*d244*(*e178*::Topo) produced is nearly 10-fold lower, however, suggesting that the defect caused by the *d244* deletion results in a decreased amount of virus production per cell.

It is likely that a synchronous infection was not obtained due to low adsorption of SSV1. The adsorption constant for SSV9 is reported to be quite low, $\sim 10^{-11}$ $mL^{-1}min^{-1}$ (Bautista et al., 2015). Data for the SSV1 adsorption constant has not been published, however, data from an undergraduate honors thesis reported that SSV1 at an initial titer of 10^5 HFU/mL (halo-forming units/mL) became undetectable within 5 hours following incubation in *S. solfataricus* G-theta at 23º C. The MOI of this incubation was 0.01 (Drummond 2010). It is possibly relevant that these data used a different host and that the method of

detection was a halo assay (HFU/mL) which is less sensitive than the plaque assay (data not shown). The adsorption constant for SSV1 to S441 remains to be determined.

The latent period corresponds to the amount of time that elapses following virus adsorption and the appearance of new infectious virus in the culture. This value has not been determined for SSV1 or any fusellovirus due to the lack of one-step growth curves. Each of the one-step growth curves displays an increase in the viral titer during the first 12 h.p.i., indicating that this time frame may roughly correspond to the latent period (Figure 5-5). However, because titers were not measured at earlier time points, the exact timing of the latent period cannot be determined at this time. Because the transcription cycle after UV-irradiation requires ~8.5 hr to be completed, infectious virus should not be produced prior to this time (Fröls et al., 2007). Thus we hypothesize that the latent period is somewhere between 8.5 and 12 hours. This is supported by the low MOI infection with SSV1, where the viral titer did not increase within the first 8 h.p.i. (Figure 5-2B).

Cell growth during SSV1 and SSV1(e178::Topo) one-step growth curves

The change in MOI resulted in dramatic differences in the growth of SSV1- and SSV1(*e178*::Topo)-infected cultures. Cultures infected at a high MOI exhibited a rapid reduction in growth (<12 h.p.i.) that was strikingly similar to what has been observed in SSV9 infections at a low MOI (Bautista et al., 2015; Figure

5-3). While a reduction in the host growth rate was observed in low MOI infections, the onset was later (~ 24 h.p.i.) and not as dramatic as in high MOI infections. The reduction in host growth appears to be correlated with the abundance of virus, as low MOI cultures of SSV1 and SSV1(*e178*::Topo) received nearly 50 times lower dose of virus. Only once the titer increased (~24 h.p.i.) did a noticeable reduction in cell growth occur. It is likely that the decrease in host growth is caused by the stress imposed on the cells by viral production. Our high MOI experiments illustrated that virus is clearly produced within 12 h.p.i. (Figure 5-5). Newly produced virus in low MOI infections could then infect the remaining uninfected cells, which in turn begin to produce virus and result in the observed decrease in the average rate of cell growth.

A similar reduction in host growth was observed in SSV2-infected *S. solfataricus* P2 cells (Contursi et al., 2006). An SSV2-infected culture of P2 was established and found to exhibit an extremely low rate of growth relative to uninfected controls. However, single colonies isolated from two rounds of streaking and colony picking were found to grow at rates indistinguishable from the uninfected control, despite still being infected with SSV2. The authors hypothesized that once SSV2 establishes a stable carrier state, the host is able to resume normal growth. Following infection of S441 at high MOI with SSV1(*e178*::Topo), we see a similar reduction in growth. We also observed that washing the SSV1(*e178*::Topo)-infected culture and re-suspending the cells in fresh YS medium resulted in an improved rate of growth, although growth of

these cells was still impaired compared to uninfected controls (data not shown). This suggests that SSV1, like SSV2, may gradually establish a stable carrier state in host cultures. Alternatively, the increased rate of growth we observe could be due to resistant cells within the population that begin to dominate late in growth, as has been seen in studies with other *Sulfolobus* viruses (Bize et al., 2009).

SSV1Δd244(e178::Topo)*-infected cells display a distinct phenotype*

The growth of SSV1Δ*d244*(e178::Topo)-infected cultures (Figure 5-5B) appears dramatically different than cultures infected with SSV1 or SSV1(*e178*::Topo). At ~12 h.p.i., SSV1Δ*d244*(e178::Topo)-infected cells appear to grow at an increased rate relative to the uninfected controls (Figure 5-5B). This increased growth rate was observed in each of three independent experiments (Figure 5-5B). It is possible that the measured increase in cell density may not be due to increased cell division but may have been caused by an increase in cell size. Infected cells were examined under the light microscope, but we did not identify any obvious differences relative to uninfected controls. Follow-up experiments measuring total CFU's should be performed, perhaps in combination with flow cytometry (Okutan et al., 2013; Bernander et al., 1997).

Between 24 – 40 h.p.i., the growth rate of the SSV1Δ*d244*(e178::Topo)-infected culture began to decrease and a reduction in cell density was observed

along with the appearance of cell debris. Interestingly, the cell density did not decrease in any of the infections with SSV1, SSV1(e178::Topo), or SSV9. It is unclear why deletion of the *d244* gene from SSV1 results in the aberrant growth curve shown here and why cell death seems to occur earlier compared to other high MOI infections.

Our data indicate that SSV1Δ*d244*(e178::Topo) virus elicits a dramatic cellular response following infection at high MOI. Furthermore, this mutant resulted in a ten-fold decrease in virus production compared to wild type SSV1 and SSV1(*e178*::Topo) (Table 5-3). The function of D244 in the SSV1 life cycle is unknown. Mass spectrometry of purified SSV1 virions indicated D244 is packaged within the virion, although D244 was not identified in more highly purified virions in a more recent analysis (Menon et al., 2008; Quemin et al., 2015). The crystal structure of D212, a D244 homologue in *Sulfolobus* spindle-shaped virus 8, was solved and shares significant homology with members of the PD-(D/E)XK nuclease superfamily (Menon et al., 2008). The *d244* gene is moderately conserved in the *Fuselloviridae* and is apparently non-essential for SSV1 infectivity (Chapters 2 and 3). Furthermore, *S. solfataricus* G-theta cells infected with Δ*d244* virus exhibited a retarded growth phenotype relative to cells infected with wild-type virus. However, the previous work was performed using a culture that was not synchronously infected (Chapter 2).

High initial starting titers in one-step growth curves

It has not escaped our notice that the initial titers in the one-step growth curves are quite high (~10^5 PFU/mL). This indicates that the wash steps were inefficient and did not result in a complete removal of virus. Following the four-hour incubation with virus, cells were washed two times prior to the start of the growth curve experiments. As mentioned above, SSV1 forms aggregates with other virions and vesicles due to interactions facilitated by the tail fibers (Quemin et al., 2015). It is possible that these aggregates pellet with the cells during the wash steps, preventing their removal. These aggregates may be detected by the plaque assay as they essentially become fixed in place by the soft-layer, facilitating adsorption to any cells that are in proximity. However, this remains to be investigated.

Appearance of cell debris in high MOI infections

It is notable that cell debris was observed within 60 h.p.i. in all of the cultures infected at high MOI, regardless of the SSV being used. The presence of cell debris implies cell death, but it is unclear if this is due to lysis by SSV1 because SSV1 is thought to be non-lytic (Martin et al., 1984; Quemin et al., 2015). Work with SSV9 suggested that the mere presence of non-infectious virus was enough to trigger a state of dormancy in the population, and if dormancy persists the cells eventually die (Bautista et al., 2015). It is possible that the cell death observed here may also be due to an extended dormant state

and not by lysis caused by virus egress. We did not observe the same dormancy with UV-inactivated virus as observed by Bautista et al. (2015) making this unlikely. However, two lytic crenarchaeal viruses (SIRV2 and STIV1) were initially thought to be non-lytic (Bize et al., 2009; Ortmann et al., 2008). Therefore, it is possible that SSV1 is lytic or that SSV1 infection at high MOI leads to cell death.

Infection with SSV9 at low MOI

As mentioned above, *S. islandicus* strain RJW002 incubated with SSV9 virions at an MOI of 0.01 results in a population-wide delay in growth, where cells were viable but not growing (Bautista et al., 2015). Within 24 hours, these cells recover and resume a rate of growth comparable to uninfected controls without any SSV9 being produced. Bautista and colleagues proposed that the cells enter into a dormant state, hypothesized to provide cells time to mount an immune response (via the CRISPR-Cas pathway) to eliminate the invading phage/virus (Makarova et al., 2012). RJW002 cells with a deactivated CRISPR-Cas system that were challenged with SSV9 resulted in production of new SSV9 and a more prolonged cell dormancy. This led to the hypothesis that cell stasis in RJW002 was caused by the continuous presence of SSV9 in the culture. To test this, UV-inactivated SSV9 was mixed with RJW002 cells at a MOI of 0.01 (based on titer prior to UV-inactivation). Interestingly, under these conditions RJW002 cells still appear to enter into a dormant state from which they do not recover. Because a

dormant-like state was induced at low MOIs (i.e. less than one virus per cell) and in the presence of non-infectious virions, it was proposed that this is an anti-viral response that is independent of infection.

Incubation of S441 with SSV9 at a low MOI (0.1) displays some similarities with the SSV9-susceptible RJW002 mutants from Bautista et al. (2015). Shortly after infection with SSV9 (<12 h.p.i.), S441 cells display a pronounced decrease in growth from which they do not recover (Figure 5-3). The OD_{600nm} of the SSV9-infected culture increased by only 0.033 units between 12 and 46 h.p.i., compared to 0.767 units for the uninfected control (Figure 5-3). Our experiments display a burst in SSV9 titer of ~10^5 PFU/mL after 8 h.p.i., nearly two orders of magnitude higher than the 10^3 – 10^4 PFU/mL burst observed by Bautista et al. at 24 h.p.i. This burst coincides well with the onset of growth inhibition in the host culture. The minimal growth of SSV9-infected S441 cells mirrors that of SSV9-infected RJW002 cells with an inactivated CRISPR-Cas system and lends support to their hypothesis that the reduction of cell growth (i.e. dormancy) is due to prolonged exposure to SSV9 (Bautista et al., 2015). However, this stasis-like state could simply be caused by the stress imposed on the cells by viral production, similar to what is seen in SSV1 infections at high MOI, and to a lesser extent at low MOI (Figures 5-2 and 5-5).

It is odd that low MOIs of SSV9 appear to cause a similar growth defect (in terms of time of onset) as the high MOI infections with SSV1. Perhaps infection of only a subset of the population by SSV9 does result in a “signal” to the

remaining cells, “warning” them of the danger as previously hypothesized (Bautista et al., 2015). We have found that infection with ~50 times more SSV1 (MOI = 5) is required to elicit a similar cell response (Figure 5-5). This amount of SSV1 is theoretically equal to five infectious particles per cell, suggesting that the growth defects observed could be explained by the stress of virus production and is intrinsically different than the effect caused by low MOI SSV9 infections. This is supported by the observation that low MOI (0.01) SSV1 infections do not result in a growth defect until much later.

We next investigated if UV-inactivated SSV1(*e178*::Topo) was capable of inducing a similar dormancy-like state as was previously shown for UV-inactivated SSV9 (Bautista et al., 2015). UV-inactivated SSV1(*e178*::Topo) was mixed with S441 cells at a MOI of 5, based on the viral titer calculated prior to UV-inactivation (Figure 5-6). The inactivated virus originated from the same stock that previously resulted in a severe decrease in growth of S441 (Figure 5-4). Challenge of S441 at a much higher inactive virion-to-cell ratio did not result in any signs of host growth inhibition relative to the uninfected control. Our results indicate that virions of SSV1(*e178*::Topo), and SSV1 by extension, do not affect cell growth as was observed elsewhere with SSV9 (Bautista et al., 2015). This may indicate a fundamental difference between SSV1 and SSV9. It will be interesting to see if UV-inactivated SSV9 is capable of inducing a dormancy-like effect in S441. These experiments will help clarify if the virion-induced stasis caused by UV-inactivated SSV9 is observed in a different host (S441), providing

further support for Bautista's hypothesis that cell dormancy is an antiviral defense mechanism that is independent of infection.

Sulfolobicins, a family of antimicrobial peptides found in *Sulfolobus*, provide an alternative explanation for the growth defect observed by Bautista et al. (2015) following challenge with UV-inactivated SSV9. Antimicrobial peptides are common in nature, but had not been identified in *Sulfolobus* until relatively recently (Prangishvili et al., 2000; Besse et al., 2015). Forty-one strains of *S. islandicus* were shown to exhibit antimicrobial activity towards *S. solfataricus* P1 and *S. shibatae*, although none of the *S. islandicus* strains were inhibited by their own sulfolobicins nor by the sulfolobicins of the other producer strains. One of these S. *islandicus* strains (HEN2/2) was analyzed more closely and a 20-kDa peptide was found associated with the cell S-layer and with vesicles released from these cells. *S. acidocaldarius* was later shown to encode two sulfolobicin genes (*SulA* and *SulB*), both of which are required for activity of the SulAB sulfolobicin. SulAB was shown to impede the growth of *S. islandicus* HEN2/2, suggesting that the sulfolobicins from these cells are different and hinting that their may be a variety of these molecules in nature (Ellen et al., 2011). Because lysis has not been demonstrated following exposure to sulfolobicins, it is hypothesized that the mechanism of action works by causing a growth defect, similar to what we observe in fusellovirus infections (Ellen et al., 2011).

It can be argued that the effects observed by Bautista et al. (2015) following challenge with UV-inactivated SSV9 could be explained by the action of

sulfolobicins. Bautista et al. (2015) purified the SSV9 virus used in all of their experiments from *S. islandicus* strain GV.10.6, a different strain from the one in which the infections were performed in (*S. islandicus* RJW002). Although these strains are closely related, it is unknown if this strain would be resistant to any sulfolobicins theoretically produced by GV.10.6. Furthermore, the negative control in these experiments consisted of spent *S. islandicus* RJW002 medium (in place of virus) and not spent medium from a virus-free culture of GV.10.6. Previous studies have shown that each strain is immune to its own sulfolobicins (as would be expected). Thus the negative control used in Bautista et al. (2015) would not reflect the activity of any sulfolobicins present in their SSV9 preparations, as RJW002 would be immune. Although the SSV9 preparations were concentrated using a 30 kDa spin-column, the sulfolobicin produced by strain HEN2/2 elutes with 30-40 kDa proteins via SEC chromatography and SulAB produced by *S. acidocaldarius* has a theoretical molecular weight of 40 kDa (Prangishvili et al., 2000; Ellen et al., 2011). The minimum sulfolobicin concentration required to elicit a growth defect is unknown as is the effect of a viral infection on the production of these compounds. It is highly unlikely that any growth defects we observed in our infections are the result of sulfolobicins as all of the virus strains used were isolated from the same host strain (S441).

Comparisons to other archaeal viruses

One-step growth curves have only been performed on two other crenarchaeal viruses, both of which are lytic. Infection of *S. islandicus* with SIRV2 at a MOI of 7 resulted in an almost immediate cessation of host growth after which lysis shortly follows (Bize et al., 2009). A follow-up study (in a different host, *S. solfataricus* P2) showed a burst size of ~30 virions/infected cell occurs after a latent period of about 12 hours, the timing of which is similar to our results with SSV1 shown here (Okutan et al., 2013). However, the burst sizes for SSV1 cannot be productively compared to SIRV2, as we probably did not achieve synchronous infections. Infection of *S. solfataricus* with STIV1 at MOIs of 1.5 – 2 resulted in a more delayed production of virus relative to SIRV2 and SSV1 (~16 - 32 h.i.p.), although this study measured intracellular viral genomes rather than infectious virus (Ortmann et al., 2008). The only one-step growth curve reported for a non-lytic archaeal virus was done with the halovirus HHPV-1 (Roine et al., 2010). Infection of *H. hispanica* at an MOI of 15 resulted in a maximum HHPV-1 titer (~10^{11} PFU/mL) at 25 h.p.i. and resulted in a similar reduction in host growth (without apparent lysis) observed in infections with other crenarchaeal viruses.

It should be noted that originally infection with SIRV2 and STIV1 resulted in only a small minority of infected cells (~10%) while the majority of the population appears resistant (Ortmann et al., 2008; Okutan et al., 2013). Separate *S. solfataricus* P2 strains were isolated that are highly susceptible to

SIRV2 or STIV1 infection. The SIRV2-susceptible strain (*S. solfataricus* 5E6) harbors a large deletion in the CRISPR-Cas loci, potentially explaining the loss of resistance to SIRV2. The genetic explanation for resistance to STIV1 is unknown. We observed a similar increase in cell density for high MOI-infected cultures at long times with all virus infections (data not shown). Therefore, it is possible that a similar phenomenon may be at work (i.e. some cells in the population are resistant). It would be worthwhile to investigate this phenomenon, as isolation of a highly susceptible host could result in improved viral titers and could also improve the accuracy of one-step growth curves and the information about SSV1 infections that they provide.

Conclusions

Little is known about the fusellovirus life cycle, let alone the vast majority of archaeal viruses. To this end, we have pursued one-step growth curves with SSV1 and SSV1 mutants as these will be essential to better understand this unique family of viruses. The work detailed in this chapter has detailed an efficient and reliable protocol for the production of large amounts of infectious SSV1 ($> 10^9$ PFU/mL) which will be useful for one-step growth curve experiments and other applications. Although this protocol was successfully applied to SSV1 mutants, it was not applicable to SSV9. It is unclear if SSV1 is an outlier or if the other fuselloviruses will be similarly tractable.

High viral titers isolated from low MOI infections were employed to perform one-step growth curve experiments at a MOI of 5. A MOI of 5 is theoretically sufficient to achieve a synchronous infection, however, the low burst sizes and high baseline titers obtained from our experiments indicate this was not achieved. The reasons for this are unclear but may involve the adsorption rate of SSV1. Clearly more work needs to be done to optimize SSV1 one-step growth curves. Although a true one-step growth curve was not successful, our experiments do allow for comparisons to be made between different fuselloviruses. It was shown here that SSV1 and the nearly identical mutant SSV1(*e178*::Topo) produce roughly equal amounts of virus and yield very similar growth curves. Interestingly, the mutant SSV1Δ*d244*(*e178*::Topo) yielded nearly ten-fold lower virus production and resulted an aberrant growth curve compared to the other viruses analyzed. We have isolated a large number of infectious SSV1 mutants, the vast number of which do not display an apparent phenotype (Chapter 3). One-step growth experiments similar to those performed here will be useful for identifying the effects of these mutations and may also help elucidate the functions of mysterious SSV1 ORFs.

SSV9, as has been detailed elsewhere (Bautista et al., 2015), causes a rapid decrease in host growth at low MOI. This defect is reminiscent of the defect caused in high MOI infections with SSV1 but was not observed in low MOI SSV1 infections until much later in the experiment. This suggests that there is a fundamental difference between these two viruses that causes the host to react

quite differently to an infection. Although non-infectious SSV9 virions apparently elicit a similar growth defect, we did not observe a similar defect using UV-inactivated SSV1(*e178*::Topo) virions. Based on this result, we have proposed an alternative explanation for the SSV9 data that involves the antimicrobial peptide sulfolobicin. Clearly more research will be needed to clarify and expand upon these interesting results.

Chapter 6: Summary and future directions

In the preceding chapters we have presented an extensive analysis of SSV1 genetics, biochemistry, and the kinetics of infection. Using both specific and random mutagenesis, we showed that the genome of SSV1 is highly malleable and encodes a large proportion of apparently non-essential ORFs, including the well conserved minor capsid gene *vp3*. A method for the purification of recombinant SSV1 structural protein VP2 was presented that should be helpful for future research aimed at better characterizing the biochemical and structural characteristics of this fusellovirus protein. Finally, we provided a technique for the isolation of large amounts of infectious SSV and used this to demonstrate the first known one-step growth curves for a fusellovirus. This work has provided insights into the fusellovirus lifecycle and should also spawn a variety of follow-up studies. A brief summary of our findings and ideas for future research are now presented.

The role of VP2 in the Fuselloviridae

Many questions about the structure and function of VP2 remain. The observation that the SSV1 genome is positively supercoiled raises questions related to how supercoiling arises and if VP2 plays any role in its establishment or maintenance (Nadal et al., 1986). To address this, it would be illuminating to investigate the DNA packaged within other fuselloviruses that do not encode a VP2 homologue. It would also be worthwhile to look at the other VP2-encoding

fuselloviruses (i.e. SSV6 and ASV1) to investigate if the topology of the packaged DNA is correlated with the presence of a VP2 homologue. We have also isolated an infectious mutant of SSV1 lacking the *vp2* gene whose packaged genome should also be investigated. A protocol for the isolation of large amounts of virus was presented in chapter 5 and could be useful for these experiments.

It is intriguing that purified SSV1 virions appear to harbor a host DNA binding protein (Sso7d) in addition to VP2 (Quemin et al., 2015). VP2 shares many characteristics with the numerous chromatin-associated proteins of *Sulfolobus* suggesting a functional overlap (Zhang et al., 2012). The crenarchaeal virus STIV1 has also been shown to harbor Sso7d, indicating this may be a common viral strategy in these environments (Maaty et al., 2006). It is unknown if other fuselloviruses adopt a similar strategy, but it seems likely. To investigate this, the structural proteins from a different fusellovirus isolate should be purified and analyzed by mass spectrometry. Not only will this provide information on what DNA binding proteins exist within the other fusellovirus capsids, it will also expand our knowledge of fusellovirus structural proteins beyond SSV1. A good candidate would be either of the well characterized viruses SSV2 or SSV9, neither of which encodes a VP2 homolog or the VP2 deletion mutant.

The protocol presented in chapter 2 for the purification of 6-His VP2 should aid future research aimed at characterizing the biochemical and structural

properties of VP2. We have shown that 6-His VP2 can be purified under denaturing conditions to yield apparently native protein capable of binding non-specifically to dsDNA. Although VP2 is known to co-purify with the capsid proteins VP1 and VP3, direct interactions between these proteins and VP2 have never been demonstrated experimentally (Reiter et al., 1987a). Experiments using the yeast-2-hybrid system are feasible and could elucidate any interactions. We have also isolated mutants harboring mutations in the *vp1* and *vp3* structural genes. It would be interesting to see if VP2 is still present in the capsids of these mutants.

One of the major motivations for purifying recombinant VP2 is to obtain a three-dimensional structure. The structure of VP2 will be vital to better understand the mode of DNA binding and determine any structural relationships with other known DNA binding proteins. Because VP2 shares many physical properties with other *Sulfolobus* NAPs, it will be interesting to see if any structural similarities are likewise shared. Although initial crystallography screens have proven unsuccessful, preliminary NMR data suggests that this is a viable alternative should crystallography be unsuccessful.

The genetic characterization of SSV1

We have presented in chapter 3 a comprehensive genetic analysis of the fusellovirus SSV1 which has greatly expanded our knowledge of the genetic requirements of this unique archaeal virus. The utility of LIPCR in the

construction and complementation of fusellovirus mutants was further demonstrated and we also show that transposon mutagenesis is an effective and efficient method to isolate SSV1 mutants. The two methods yield highly similar results with a few exceptions that are probably explained by polar effects caused by transposon insertion. Deletion mutants in these ORFs could be constructed, although this is probably not crucial at this time. Alternatively, and perhaps more realistically, is these mutants could be complemented in trans to restore infectivity. We have shown that complementation in cis of SSV1 mutants can be performed, something that has not been done for any crenarchaeal virus to our knowledge. Due to a lack of selectable markers, complementation in trans has been challenging in *Sulfolobus* although progress has been made (Zhang et al., 2013).

Interestingly, infectious virus harboring various mutations within the *vp1* and *vp3* structural genes were able to be isolated. The finding that SSV1 remains infectious despite the loss of the minor capsid gene *vp3* was quite unexpected. We hypothesize that the homology that exists between VP1 and VP3 enables the VP1 protein to rescue this mutant via complementation. However, the loss of VP3 does correlate with the appearance of elongated particles suggesting complementation is either not completely effective or simply does not occur. Higher resolution images of these particles will be required to better understand the effect of this mutation.

The major capsid protein VP1, while not completely dispensable, does tolerate mutations to the universally conserved glutamate believed to be crucial for proteolytic cleavage during the maturation of VP1 (Reiter et al., 1987a; Quemin et al., 2015). Despite its absence in the mature VP1 protein, the VP1 N-terminus could not be deleted without eliminating virus infectivity. This suggests an essential role for the VP1 N-terminus and parallels what has been observed for the MCP of bacteriophage HK97. An attractive hypothesis is that the N-terminus of VP1 may act as a scaffold (or may interact with a different scaffolding protein, e.g. A291, see below) during assembly of VP1, a role which has not been assigned to any SSV1 protein. The N-terminus of VP1 may also be important for recruitment of the viral protease, which has not been identified. The recent observation that the product of ORF *b251* is homologous to lon proteases is intriguing and presents a good candidate for future studies (Happonen et al., 2014). Our mutagenesis data indicate this gene is essential for SSV1 infectivity, lending support to this hypothesis.

The abundance of T3 and Tx transcripts (encoding ORFs *a291* and *c124*, respectively) in stably infected cells, along with expression of structural genes suggests a potential structural role for SSV1-A291 and/or SSV1-C124 (Fusco et al., 2015c). SSV1-A291, and to a lesser extent SSV1-C124, are conserved in the *Fuselloviridae* supporting this conclusion. The well conserved N-terminus of SSV1-A291 could potentially interact with other structural proteins during assembly in scaffolding role. Yeast-two-hybrid screens to detect interactions

between A291 and the other SSV1 capsid proteins (vp1, vp2, vp3, and vp4) should be fairly straightforward.

The effect of mutations to the conserved glutamate on VP1 maturation are unknown. It is possible that cleavage occurs at an alternate residue or that cleavage does not occur and is not essential for infectivity. It would be very worthwhile to analyze these mutants, along with Δ*vp3* virus, via mass spectrometry. Furthermore, the effect of these structural mutations on the thermal stability of the SSV1 capsid should be investigated in addition to one-step growth curves.

Recently it has been shown that the structural proteins VP1, VP3 and VP4 are glycosylated in the SSV1 virion (Quemin et al., 2015). Glycosylation may be critical for virion stability and/or for interactions with the host cell during adsorption. The putative sites of glycosylation have been identified, with VP1 and VP3 each containing two sites and VP4 containing a total of 20 sites. The physiological relevance of this glycosylation remains to be investigated. We have shown here that LIPCR can be employed to generate single base pair changes in the SSV1 genome. It is would be worthwhile to construct mutants with defects in these putative glycosylation sites to help comprehend their importance in the *Fuselloviridae*.

Although a significant amount of work has been done with SSV1 genetics, the other fuselloviruses have been almost completely neglected. We have shown that transposon mutagenesis can be easily applied to fuselloviruses other

than SSV1 (appendix A). Although SSV9 seems to be fairly intolerant of transposon mutagenesis, our lab has begun work with a fusellovirus isolated from Lassen National Volcanic Park (SSVL) and preliminary results indicate this virus may be similar to SSV1 in regards to its ability to tolerate insertions. Shuttle vectors based on the other fuselloviruses will open the door for construction of deletion mutants, further expanding our knowledge of fusellovirus genetics. One of the major surprises from this work was the finding that the *vp3* gene was not essential for infectivity. It will be interesting to see if the *vp3* gene in other fuselloviruses can also be removed without a loss of infectivity or if this is an SSV1-specific phenomenon. LIPCR can be performed with relative ease so these mutants should be fairly straightforward to construct.

A large number of the mutants examined in this work yielded non-infectious phenotypes. Experiments were repeated numerous times ($n \geq 5$) using rigorous controls to ensure that transformations were successful and viral DNA was transported into the cell. Restriction endonuclease digestion and sequencing of the mutated region also helped to ensure viral genomes were full length and intact. However, small deletions may have escaped our notice and could contribute to the non-infectious phenotypes observed. To address this, we plan to sequence the complete genomes of our non-infectious mutants to probe for any secondary mutations.

The kinetics of a fusellovirus infection

The first insights into the kinetics of a fusellovirus infection were presented in chapter 5. One-step growth curve experiments are useful to determine key attributes of a viral life cycle (Ellis and Delbruck 1939). Although SSV1 has been well studied, no such experiments have been performed. This is likely due to the difficulties associated with isolating enough infectious virus to perform the synchronous infection required in a one-step growth curve. We found that infection of *S. solfataricus* strain S441 at a low MOI results in the production of high titers that can then be used for high MOI infections. This protocol was successfully applied to wild type SSV1 and two SSV1 mutants, indicating it is reliable and repeatable. However, high titers of the virus SSV9 were not able to be obtained indicating this protocol may not be suited for all fuselloviruses.

One-step growth curves indicate that SSV1 (and SSV1 mutants) has a latent period of ~8.5 – 12 hr, which correlates well with previous data on the timing of the SSV1 transcription cycle in infected cells (Fröls et al., 2008). To more accurately determine the latent period, more measurements need to be taken within the first 12 hours of infection. One-step growth curves also allowed for a rough estimation of SSV1 burst size. An apparent burst size of ~3 virions per infected cell was measured, although this may be an underestimate as we are unsure if a synchronous infection was achieved. To address this, experiments measuring the SSV1 adsorption rate to *Sulfolobus* strain S441 need

to be performed which will allow a more for a more accurate viral adsorption period during one-step growth curve experiments.

Although a synchronous infection may not have been achieved, our experiments do allow for comparisons to be made between SSV1, SSV1 mutants, and other fuselloviruses. We have isolated a number of infectious SSV1 mutants, almost all of which do not have obvious phenotypes. One-step growth curves will be useful for teasing out any differences that may exist due to these mutations. We found that SSV1 harboring a deletion of the *d244* ORF exhibits a dramatically different growth curve relative to wild type SSV1. Furthermore, infection with Δ*d244* results in a 10-fold lower burst size than wild-type virus and apparently results in an acceleration of host cell growth shortly following infection. It will be interesting to see how the growth curves of other SSV1 mutants compare.

Low MOI infections with SSV9 resulted in a significant growth defect from which the cells do not appear to recover. This is very similar to the effect caused by high MOI infections with SSV1 but differs from low MOI SSV1 infections. We also show that unlike UV-inactivated SSV9, UV-inactivated SSV1 has no effect on host growth. We postulated an alternate explanation for what has been observed by others in SSV9 based on the family of antimicrobial peptides known as Sulfolobicins that are common in *Sulfolobus*. It is unclear why SSV9 elicits a more dramatic cellular response at lower MOIs than SSV1. Very little work has been done on the other fuselloviruses, so it will be interesting to see if this is

unique to SSV9 or if this is a common phenomenon in the *Fuselloviridae* and SSV1 is the exception. Preliminary work on SSV3 and SSVL have indicated similar phenotypes to SSV9, but this work is very much in its infancy.

The protocol presented here for the isolation of high SSV1 titers has a few other applications. As detailed elsewhere in this dissertation, we have isolated a number of infectious SSV1 structural mutants and the effects of these mutations on the SSV1 virion are poorly understood (Chapter 3). Techniques to investigate the structure and composition of SSV1 virions exist (cryo-EM reconstructions, mass spectrometry) but require significant amounts of virus. The protocol presented here should be useful in these pursuits. Furthermore, the SSV1 transcription cycle has been well characterized but only following UV induction and there is evidence that the transcription cycle differs following a "natural" infection. This technique also requires a significant amount of virus that can be achieved by applying the results of the work presented herein.

Conclusion

Although a significant amount of research has been devoted to the *Fuselloviridae*, there is still much we do not know. The research presented in this dissertation has provided some significant insights into the life-cycle and genetic requirements of SSV1, helping to further characterize this unique family of archaeal viruses while at the same time opening new avenues of research. We have further honed the genetic techniques available for fuselloviruses and

provided some intriguing targets for future studies. The identification of a number of infectious structural mutants should provide valuable insight into the assembly and thermal stability of spindle-shaped capsids. A putative scaffolding role for the seemingly unimportant VP1 N-terminus has been suggested and should be investigated. The elongated virus particles produced by *vp1* and *vp3* structural mutants are intriguing and await higher-resolution imaging and proteomics to better explain these results. We have isolated a number of infectious mutants with defects in a variety of genes, the vast majority of which display no apparent phenotype. Analysis of these mutants by one-step growth curves should be straightforward and, as we have seen with one mutant already, has the potential to reveal abnormalities in the SSV infection cycle and help elucidate the function of mysterious SSV1 proteins.

References cited:

Albers, S.V., Forterre, P., Prangishvili, D., and Schleper, C. (2013). The legacy of Carl Woese and Wolfram Zillig: from phylogeny to landmark discoveries. *Nat. Rev. Microbiol*. 11 (10), 713-719.

Andersson, A.F., and Banfield, J.F. (2008). Virus population dynamics and acquired virus resistance in natural microbial communities. *Science*. 320 (5879), 1047-1050.

Arnold, H.P., She, Q., Phan, H., Stedman, K., Prangishvili, D., Holz, I., Kristjansson, J.K., Garrett, R., and Zillig, W. (1999). The genetic element pSSVx of the extremely thermophilic crenarchaeon *Sulfolobus* is a hybrid between a plasmid and a virus. *Mol. Microbiol*. 34 (2), 217-226.

Bautista, M.A., Zhang, C., and Whitaker, R.J. (2015). Virus-induced dormancy in the archaeon *Sulfolobus* islandicus. *MBio*. 6 (2).

Barrangou, R. and Marraffini, L.A. (2014). CRISPR-Cas systems: Prokaryotes upgrade to adaptive immunity. *Mol. Cell*. 54 (2), 233 -234.

Barry, E.R., and Bell, S.D. (2006). DNA replication in the Archaea. *Microbiol. Mol. Biol. Rev*. 70, 876-877.

Bell, S.D., Bottling, C.H., Wardleworth, B.N., Jackson, S.P., and White,

M.F. (2002). The interaction of Alba, a conserved archaeal chromatin protein, with Sir2 and its regulation by acetylation. *Science*. 296 (5565), 148-151.

Benson, S.D., Bamford, J.K., Bamford, D.H., and Burnett, R.M. (2004). Does common architecture reveal a viral lineage spanning all three domains of life? *Mol. Cell*. 16 (5), 673-685.

Bernander, R., and Poplawski, A. (1997). Cell cycle characteristics of the thermophilic Archaea. *J. Bacteriol.* 179 (16), 4963 - 4969.

Besse, A., Peduzzi, J., Rebuffat, S., and Carré-Mlouka, A. (2015). Antimicrobial peptides and proteins in the face of extremes: Lessons from archaeocins. *Biochimie*. Advanced online publication. doi: 10.1016/j.biochi.2015.06.004.

Birnboim, H.C., and Doly, J. (1979). A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acids Res*. 7, 1513-1523.

Bize, A., Karlsson, E.A., Ekefjärd, K., Quax, T.E., Pina, M., Prevost, M.C., Forterre, P., Tenaillon, O., Bernander, R., and Prangishvili, D. (2009). A unique virus release mechanism in the Archaea. *Proc. Natl. Acad. Sci. USA*. 106 (27), 11306-11311.

Bolduc, B., Shaughnessy, D.P., Wolf, Y.I., Koonin, E.V., Roberto, F.F., and Young, M. (2012). Identification of novel positive-strand RNA viruses by

metagenomic analysis of Archaea-dominated Yellowstone hot springs. *J. Virol*. 86 (10), 5562-5573.

Brochier-Armanet, C., and Forterre, P. (2007). Widespread distribution of archaeal reverse gyrase in thermophilic bacteria suggest a complex history of vertical inheritance and lateral gene transfers. *Archaea*. 2 (2), 83-93.

Brock, T.D., Brock, K.M., Belly, R.T., and Weiss, R.L. (1972). *Sulfolobus*: a new genus of sulphur oxidizing bacteria living at low pH and high temperature. *Arch. Mikrobiol*. 84 (1), 54-68

Brumfield, S.K., Ortmann, A.C., Ruigrok, V., Suci, P., Douglas, T., and Young, M.J. (2009). Particle assembly and ultrastructural features associated with replication of the lytic archaeal virus *Sulfolobus* turreted icosahedral virus. *J. Virol.* 83 (12), 5964-5970.

Cann, A.J. (2016). Principles of Molecular Virology. London (England): Academic Press. Chapter 3: The structure and complexity of virus genomes.

Cannio, R., Contursi, P., Rossi, M., and Bartolucci S. (1998). An autonomously replicating transforming vector for *Sulfolobus solfataricus*. *J Bacteriol*. 180 (12), 3237-3240.

Cavicchioli, R. (2011). Archaea — timeline of the third domain. *Nat. Rev. Mircobiol.* 9 (1), 51 - 61.

Ceballos, R.M., Marceau, C.D., Marceau, J.O., Morris, S., Clore, A.J., and Stedman, K.M. (2012). Differential virus host-ranges of the *Fuselloviridae* of hyperthermophilic Archaea: implications for evolution in extreme environments. *Front. Microbio.* doi: 10.3389/fmicb.2012.00295.

Clore, A.J., and Stedman, K.M. (2006). The SSV1 viral integrase is not essential. *Virology*. 361 (1), 103-111.

Clore, A.J. (2008). The family *Fuselloviridae*: Diversity and replication of a hyperthermophilic virus infecting the archaeon genus *Sulfolobus*. (Doctoral dissertation, Portland State University).

Contursi, P., Jensen, S., Aucelli, T., Rossi, M., Bartolucci, S., and She, Q. (2006). Characterization of the *Sulfolobus* host-SSV2 virus interaction. *Extremophiles*. 10 (6), 615-627.

Contursi,P., Fusco, S., Cannio, R., and She, Q. (2014). Molecular biology of fuselloviruses and their satellites. *Extremophiles*. 18 (3), 473-489.

Dellas, N., Lawrence, C.M., and Young, M.J. (2013). A survey of protein structures from archaeal viruses. *Life (Basel).* 3(1), 118-130.

Delong, E.F. (1998). Everything in moderation: Archaea as 'non-extremophiles'. *Curr. Opin. Genet. Dev.* 8 (6), 649-654.

Dougherty, W.G., and Semler, B.L. (1993). Expression of virus-encoded proteinases: functional and structural similarities with cellular enzymes. *Microbiol. Rev.* 57 (4), 781-822.

Driessen, R.P., and Dame, R.T. (2011). Nucleoid-associated proteins in Crenarchaea. *Biochem. Soc. Trans*. 39 (1), 116-121.

Driessen, R.P., Meng, H., Suresh, G., Shahapure, R., Lanzani, G., Priyakumar, U.D., White, M.F., Schiessel, H., van Noort, J., and Dame, R.T. (2013). Crenarchaeal chromatin proteins Cren7 and Sul7 compact DNA by inducing rigid bends. *Nucleic Acids Res.* 41(1), 196-205.

Drozdetskiy, A., Cole, C., Procter, J., and Barton, G.J. (2015). JPred4: a protein secondary structure prediction server. *Nucleic Acids Res.* 143 (W1), W389 - 394.

Drummond, C. (2010). Characterization of *Sulfolobus* spindle-shaped virus replication. (Undergraduate honors thesis, Portland State University).

Duda, R.L., Hempel, J., Michel, H., Shabanowitz, J., Hunt, D., and Hendrix, R.W. (1995a). Structural transitions during bacteriophage HK97 head assembly. *J. Mol. Biol*. 247, 618-635.

Duda, R.L., Martincic, K., and Hendrix, R.W. (1995b). Genetic basis of

bacteriophage HK97 prohead assembly. *J. Mol. Biol*. 247, 636-647.

Eilers, B.J., Young, M.J., and Lawrence, C.M. (2012). The structure of an archaeal viral integrase reveals an evolutionarily conserved catalytic core yet supports a mechanism of DNA cleavage in trans. *J. Virol*. 86(15), 8309-8313.

Ellen, A.F., Rohulya, O.V., Fusetti, F., Wagner, M., Albers, S.V., and Driessen, A.J. (2011). The sulfolobicin genes of *Sulfolobus* acidocaldarius encode novel antimicrobial proteins. *J Bacteriol.* 193 (17), 4380 - 4387.

Ellis, E.L., and Delbrück, M. (1939). The growth of bacteriophage. *J. Gen. Physiol.* 22 (3), 365-384.

Erdmann, S. and Garrett, R.A. (2012). Selective and hyperactive uptake of foreign DNA by adaptive immune systems of an archaeon via two distinct mechanisms. *Mol. Microbiol*. 85 (6), 1044-1056.

Erdmann, S., Shah, S.A., and Garrett, R.A. (2013). SMV1 virus-induced CRISPR spacer acquisition from the conjugative plasmid pMGB1 in *Sulfolobus solfataricus* P2. *Biochem. Soc. Trans*. 41 (6), 1449-1458.

Erdmann, S., Le Moine Bauer S., and Garrett, R.A. (2014a). Inter-viral conflicts that exploit host CRISPR immune systems of *Sulfolobus*. *Mol. Microbiol.* 91 (5), 900-917.

Erdmann, S., Chen, B., Huang, X., Dengl, L., Liu, C., Shah, S.A., Le Moine Bauer, S., Sobrino, C.L., Wang, H., Wei, Y., She, Q., Garrett, R.A., Huang, L., and Lin, L. (2014b). A novel single-tailed fusiform *Sulfolobus* virus STSV2 infecting model *Sulfolobus* species. *Extremophiles*. 18 (1), 51-60.

Fröls, S., Gordon, P.M., Panlilio, M.A., Schleper, C., and Sensen, C.W. (2007). Elucidating the transcription cycle of the UV-inducible hyperthermophilic archaeal virus SSV1 by DNA microarrays. *Virology*. 365(1), 48-59.

Fusco, S., She, Q., Bartolucci, S., and Contursi, P. (2013). T(lys), a newly identified *Sulfolobus* spindle-shaped virus 1 transcript expressed in the lysogenic state, encodes a DNA-binding protein interacting at the promoters of the early genes. *J. Virol.* 87(10), 5926-5936.

Fusco, S., Aulitto, M., Bartolucci, S., and Contursi, P. (2015a). A standardized protocol for the UV induction of *Sulfolobus* spindle-shaped virus 1. *Extremophiles*. 19 (2), 539-546.

Fusco, S., She, Q., Florentino, G., Bartolucci, S., and Contursi, P. (2015b). Unravelling the role of the F55 regulator in the transition from lysogeny to UV induction of *Sulfolobus* spindle-shaped virus 1. *J. Virol*. 89 (12), 6453-6461.

Fusco, S., Liguori, R., Limauro, D., Bartolucci, S., She, Q., and

Contursi, P. (2015c). Transcriptome analysis of *Sulfolobus solfataricus* infected with two related fuselloviruses reveals novel insights into the regulation of CRISPR-Cas system. *Biochimie*. Epub ahead of print.

Gardner, A.F., Bell, S.D., White, M.F., Prangishvili, D., and Krupovic, M. (2014). Protein-protein interactions leading to the recruitment of the host DNA sliding clamp by the hyperthermophilic *Sulfolobus* islandicus rod-shaped virus 2. *J. Virol.* 88 (12), 7105 - 7108.

Garrett, R.A., and Klenk, H. (2008). Archaea: Evolution, Physiology, and Molecular Biology. Malden (MA): Blackwell Publishing. Chapter 6: Families of DNA viruses infecting hyperthermophilic Crenarchaea.

Geslin, C., Gaillard, M., Flament, D., Rouault, K., Le Romancer, M., Prieur, D., and Erauso, G. (2007). Analysis of the first genome of a hyperthermophilic marine virus-like particle, PAV1, isolated from Pyrococcus abyss. *J. Bateriol.* 189 (12), 4510-4519.

Goulet, A., Spinelli, S., Blangy, S., van Tilbeurgh, H., Leulliot, N., Basta, T., Prangishvili, D., Cambillau, C., and Campanacci, V. (2009). The thermo- and acido-stable ORF-99 from the archaeal virus AFV1. *Protein. Sci.* 18(6), 1316-1320.

Grissa I., Vergnaud, G., and Pourcel, C. (2007). The CRISPRdb

database and tools to display CRISPRs and to generate dictionaries of spacers and repeats. *BMC Bioinformatics*. 8, 172.

Guillière, F., Danioux, C., Jaubert, C., Desnoues, N., Delepierre, M., Prangishvili, D., Sezonov, G., and Guijarro, J.I. (2013). Solution structure of an archaeal DNA binding protein with an eukaryotic zink finger fold. *PLoS One*. 8(1), e52908.

Guo, L., Feng, Y., Zhang, Z., Yao, H., Luo, Y., Wang, J., and Huang, L. (2008). Biochemical and structural characterization of Cren7, a novel chromatin protein conserved among Crenarchaea. *Nucleic Acids Res*. 36 (4), 1129-1137.

Happonen, L.J., Erdmann, S., Garrett, R.A., and Butcher, S.J. (2014). Adenosine triphosphatases of thermophilic archaeal double-stranded DNA viruses. *Cell Biosci*. 4, 37.

Hatfull, G.F. (2015). Dark matter of the biosphere: the amazing world of bacteriophage diversity. *J. Virol*. 89 (16), 8107-8110.

Häring, M., Vestegaard, G., Rachel, R., Chen, L., Garrett, R.A., and Prangishvili, D. (2005). Virology: independent virus development outside a host. *Nature*. 436 (7054), 1101-1102.

He, F., Chen, L., and Peng, X. (2014). First experimental evidence for the presence of a CRISPR toxin in *Sulfolobus*. *J. Mol. Biol*. 426 (22), 3683-3688.

Held, N.L. and Whitaker, R.J. (2009). Viral biogeography revealed by signatures in *Sulfolobus* islandicus genomes. *Environ. Microbiol*. 11 (2), 457-466.

Hendrix, R.W., Smith, M.C., Burns, R.N., Ford, M.E., and Hatfull, G.F. (1999). Evolutionary relationships among diverse bacteriophages and prophages: all the world's a phage. *Proc. Natl. Acad. Sci. USA*. 96 (5), 2192-2197.

Iverson, E., and Stedman, K.M. (2012). A genetic study of SSV1, the prototypical fusellovirus. *Front. Microbio*. 3:200 doi: 10.3389/fmicb2012.00200.

Jäälinoja, HT., Roine, E., Laurinmäki, P., Kivela, H.M., Bamford, D.H., and Butcher, S.J. (2008). Structure and host-cell intereaction of SH1, a membrane-containing, halophilic euryarchaeal virus. *Proc. Natl. Acad. Sci. USA*. 105(23), 8008-8013.

Jonuscheit, M., Martusewitsch, E., Stedman, K.M., and Schleper, C. (2003). A reporter gene system for the hyperthermophilic archaeon *Sulfolobus solfataricus* based on a selectable and integrative shuttle vector. *Mol. Microbiol*. 48 (5), 1241-1252.

Keller, J., Leulliot, N., Cambilau, C., Campanacci, V., Porciero, S.,

Prangishvili, D., Forterre, P., Cortez, D., Quevillon-Cheruel, S., and van Tilbeurgh, H. (2007). Crystal structure of AFV3-109, a highly conserved protein from crenarchaeal viruses. *Virol. J.* 4, 12.

Kessler, A., Brinkman, A.B., van der Oost, J., and Prangishvili, D. (2004). Transcription of the rod-shaped viruses SIRV1 and SIRV2 of the hyperthermophilic archaeon *Sulfolobus*. *J. Bacteriol.* 186 (22), 7745 - 7753.

Khayat, R., Tang, L., Lawrence, C.M., Young, M., and Johnson, J.E. (2005). Structure of an archaeal virus capsid protein reveals a common ancestry to eukaryotic and bacterial viruses. *Proc. Natl. Acad. Sci*. USA. 102(52), 18944-18949.

Koonin, E. (1992). Archaebacterial virus SSV1 encodes a putative DnaA-like protein. *Nucleic Acids Res*. 20 (5), 1143.

Kraft, P., Oeckinghaus, A., Kümmel, D., Gauss, G.H., Gilmore, J., Wiedenheft, B., Young, M.J., and Lawrence, C.M. (2004a). Crystal structure of F-93 from *Sulfolobus* spindle-shaped virus 1, a winged-helix DNA binding protein. *J. Virol.* 78(21) 11544-11550.

Kraft, P., Kümmel, D., Oeckinghaus, A., Gauss, G.H., Wiedenheft, B., Young, M.J., and Lawrence, C.M. (2004b). Structure of D-63 from *Sulfolobus* spindle-shaped virus 1: surface properties of the dimeric four-helix bundle

suggest an adaptor protein function. *J. virol*. 78(14) 7438-7442.

Krupovic, M., Prangishvili, D., Hendrix, R.W., and Bamford, D.H. (2011). Genomics of bacterial and archaeal viruses: dynamics within the prokaryotic virosphere. *Microbiol. Mol. Biol. Rev.* 75 (4), 610-635.

Krupovic, M., Quemin, E.R., Bamford, D.H., Forterre, P., and Prangishvili, D. (2014). Unification of the globally distributed spindle-shaped viruses of the Archaea. *J. Virol.* 88 (4), 2354-2358.

Larson, E.T., Eilers, B., Menon, S., Reiter, D., Ortmann, A., Young, M.J., and Lawrence, C.M. (2007a). A winged-helix protein from *Sulfolobus* turreted icosahedral virus points toward stabilizing disulfide bonds in the intracellular proteins of a hyperthermophilic virus. *Virology*. 368(2), 249-261.

Larson, E.T., Eilers, B.J., Reiter, D., Ortmann, A.C., Young, M.J., and Lawrence, C.M. (2007b). A new DNA binding protein highly conserved in diverse crenarchaeal viruses. *Virology*. 363(2), 387-396.

Lawrence, C.M., Menon, S., Eilers, B.J., Bothner, B., Khayat, R., Douglas, T., and Young, M.J. (2009). Structural and functional studies of archaeal viruses. *J. Biol. Chem*. 284, 12599-12603.

Lipps, G. (2006). Plasmids and viruses of the thermoacidophilic crenarchaeote *Sulfolobus*. *Extremophiles*. 10 (1), 17-28.

Liu, L., and Huang, L. (2002). Induction of the *Sulfolobus shibatae* virus SSV1 DNA replication by mitomycin C. *Chinese Science Bulletin*. 47 (11), 923-927.

Luk, A.W., Williams, T.J., Erdmann, S., Papke, R.T., and Cavicchioli, R. (2014). Viruses of Haloarchaea. *Life(Basel).* 4 (4), 681-715.

Lulchev, P., and Klostermeier, D. (2014). Reverse gyrase — recent advances and current mechanistic understanding of positive DNA supercoiling. *Nucleic Acids Res*. 42 (13), 8200-8213.

Maaty, W.S., Ortmann, A.C., Dlakić, M., Schulstad, K., Hilmer, J.K., Liepold, L., Wiedenheft, B., Khayat, R., Douglas, T., Young, M.J., and Bothner, B. (2006). Characterization of the archaeal thermophile *Sulfolobus* turreted icosahedral virus validates an evolutionary link among double-stranded DNA viruses from all domains of life. *J. Virol*. 80 (15), 7625-7635.

Maaty, W.S., Steffens, J.D., Heinemann, J., Ortmann, A.C., Reeves, B.D., Biswas, S.K., Dratz, E.A., Grieco, P.A., Young, M.J., and Bothner, B. (2012). Global analysis of viral infection in an archaeal model system. *Front Microbiol*. 3, 411.

Makarova, K.S., Anantharaman, V., Aravind, L., and Koonin, E.V. (2012). Live virus-free or die: coupling of antivirus immunity and programmed

suicide in prokaryotes. *Bio. Direct*. 7, 40.

Makarova, K.S., and Koonin, E.V. (2013). Archaeology of eukaryotic DNA replication. *Cold Spring Harb. Prospect. Med.* 3 (10), a012963.

Manica, A., Zebec, Z., Teichmann, D., and Schleper, C. (2011). In vivo activity of CRISPR-mediated virus defense in a hyperthermophilic archaean. *Mol. Microbiol*. 80 (2), 481-491.

Manica, A. and Schleper, C. (2013). CRISPR-mediated defense mechanisms in the hyperthermophilic archaeal genus *Sulfolobus*. *RNA Biol*. 10 (5), 671-678.

Marinelli, L.J., Piuri, M., Swigonova, Z., Balachandran, A., Oldfield, L.M., van Kessel, J.C., and Hatfull, G.F. (2008). BRED: a simple and powerful tool for constructing mutant and recombinant bacteriophage genomes. *PLoS One*. 3 (12), e3957.

Martin, A., Yeats, S., Janekovic, D., Reiter, W.D., Aicher, W., and Zillig, W. (1984). SAV1, a temperate u.v.-inducible DNA virus-like particle from the archaebacterium *Sulfolobus* acidocaldarius isolate B12. *EMBO J.* 3 (9), 2165-2168.

Marguet, E., and Forterre, P. (1994). DNA stability at temperatures typical for hyperthermophiles. *Nucleic Acids Res.* 22 (9), 1681-1686.

Menon, S.K., Maaty, W.S., Corn, G.J., Kwok, S.C., Eilers, B.J., Kraft, P., Gillitzer, E., Young, M.J., Bothner, B., and Lawrence, C.M. (2008). Cysteine usage in *Sulfolobus* spindle-shaped virus 1 and extension to hyperthermophilic viruses in general. *Virology.* 376(2), 270-278.

Menon, S.K., Eilers, B.J., Young, M.J., and Lawrence, C.M. (2010). The crystal structure of D212 from *Sulfolobus* spindle-shaped virus ragged hills reveals a new member of the PD-(D/E)XK nuclease superfamily. *J. Virol.* 84(12), 5890-5897.

Mochizuki, T., Krupovic, M., Pehau-Arnaudet, G., Sako, Y., Forterre, P., and Prangishvili, D. (2012). Archaeal virus with exceptional virion architecture and the largest single-stranded DNA genome. *Proc. Natl. Acad. Sci. USA*. 109 (33), 13386 - 13391.

Muskhelishvili, G., Palm, P., and Zillig, W. (1993). SSV1-encoded site-specific recombination system in *Sulfolobus shibatae*. *Mol. Gen. Genet.* 237 (3), 334 - 342.

Nadal, M., Mirambeau, G., Forterre, P., Reiter, W.D., and Duguet, M. (1986). Positively supercoiled DNA in a virus-like particle of an archaebacterium. *Nature*. 321, 256-258.

Napoli, A., Kvaratskelia, M., White, M.F., Rossi, M., and Ciaramella, M.

(2001). A novel member of the bacterial-archaeal regulator family is a nonspecific dna-binding protein and induces positive supercoiling. *J. Biol. Chem.* 276 (14), 10745-10752.

Oh, B., Moyer, C.L., Hendrix, R.W., and Duda, R.L. (2014). The delta domain of the HK97 major capsid protein is essential for assembly. *Virology*. 0, 171-178.

Ortmann, A.C., Brumfield, S.K., Walther, J., McInneerney, K., Brouns, S.J., van de Werken, H.J., Bothner, B., Douglas, T., van de Oost, J., and Young, M.J. (2008). Transcriptome analysis of infection of the archaeon *Sulfolobus solfataricus* with *Sulfolobus* turreted icosahedral virus. *J. Virol.* 82 (10), 4874-4883.

Oke, M., Kerou, M., Liu, H., Peng, X., Garrett, R.A., Prangishvili, D., Naismith, J.H., White, M.F. (2011). A dimeric Rep protein initiates replication of a linear archaeal virus genome: implications for the Rep mechanism and viral replication. *J. Virol.* 85 (2), 925 - 931.

Okutan, E., Deng, L., Mirlashari, S., Uldahl, K., Halim, M., Liu, C., Garrett, R.A., She, Q., and Peng, X. (2013). Novel insights into gene regulation of the rudivirus SIRV2 infecting *Sulfolobus* cells. *RNA Biol.* 10 (5), 875-885.

Palm, P., Schleper, C., Grampp, B., Yeats, S., McWilliam, P., Reiter, W.D., and Zillig, W. (1991). Complete nucleotide sequence of the virus SSV1 of the archaebacterium *Sulfolobus* shibatae. *Virology*. 185 (1), 242-250.

Peng, X., Blum, H., She, Q., Mallok, S., Brügger, K., Garrett, R.A., Zillig, W., and Prangishvili, D. (2001). Sequences and replication of genomes of the Archaea rudiviruses SIRV1 and SIRV2: relationships to the Archaea lipothrixvirus SIFV and some eukarayal viruses. *Virology*. 291 (2), 226 - 234.

Peng, X., Basta, T., Häring, M., Garrett, R.A., and Prangishvili, D. (2007). Genome of the Acidianus bottle-shaped virus and insights into the replication and packaging mechanisms. *Virology*. 364 (1), 237 - 243.

Peng, X. (2008). Evidence for the horizontal transfer of an integrase gene from a fusellovirus to a pRN-like plasmid within a single strain of *Sulfolobus* and the implications for plasmid survival. *Microbiology*. 154 (Pt2), 383-391.

Perez-Espino, D., Sharon, I., Morovic, W., Stahl, B., Thomas, B.C., Barrangou, R., and Banfield, J.F. (2015). *MBio*. 6 (2). pii: e00262-15. doi: 10.1128/mBio.00262-15.

Pietillä, M.K., Roine, E., Paulin, L., Kalkkinen, N., and Bamford, D.H. (2009). An ssDNAvirus infecting Archaea: a new lineage of viruses with membrane envelope. *Mol. Microbiol.* 72 (2), 307 - 319.

Pietillä, M.K., Laurinmäki, P., Russell, D.A., Ko, C.C., Jacobs-Sera, D., Hendrix, R.W., Bamford, D.H., and Butcher, S.J. (2013). Structure of the archaeal head-tailed virus HSTV-1 completes the HK97 fold story. *Proc. Natl. Acad. Sci. USA*. 110 (26), 10604 - 10609.

Prangishvili, D., Arnold., H.P., Gotz, D., Ziese, U., Holz, I., Kristjanson, J.K., and Zillig, W. (1999). A novel virus family, the Rudiviridae: structure, virus-host interactions and genome variability of the *Sulfolobus* viruses SIRV1 and SIRV2. *Genetics*. 152 (4), 1387-1396.

Prangishvili, D., Holz, I., Stieger, E., Nickell, S., Kristjansson, J.K., and Zillig, W. (2000). Sulfolobicins, specific proteinaceous toxins produced by strains of the extremely thermophilic Archaea genus *Sulfolobus*. *J. Bacteriol.* 182 (10), 2985 - 2988.

Prangishvili, D., Garrett, R.A., and Koonin, E.V. (2006). Evolutionary genomics of archaeal viruses: unique viral genomes in the third domain of life. *Virus Res.* 117(1), 52-67.

Prangishvili, D. (2013). The Wonderful World of Archaeal Viruses. *Ann. Rev. Microbiol*. 67, 565-585.

Quax, T.E., Voet, M., Sismeiro, O., Dillies, M.A., Jagla, B., Coppée, J.Y., Sezonov, G., Forterre, P., van der Oost, J., Lavigne, R., and Prangishvili, D.

(2013). Massive activation of archaeal efense genes during viral infection. *J. Virol*. 87 (15), 8419-8428.

Quemin, E.R., Pietillä, M.K., Oksanen, H.M., Forterre, P., Rijpstra, W.I., Schouten, S., Bamford, D.H., Prangishvili, D., and Krupovic, M. (2015). *Sulfolobus* spindle-shpaed virus 1 contains glycosylated capsid proteins, a cellular chromatin protein and host-derived lipids. *J. Virol*. 89 (22), 11681 – 11691.

Qureshi, SA. (2007). Protein-DNA interactions at the *Sulfolobus* spindle-shaped virus-1 (SSV1) T5 and T6 promoters. *Can. J. Microbiol.* 53 (9), 1076 - 1083.

Redder, P., Peng, X., Brügger, K., Shah, S.A., Roesch, F., Greve, B., She, Q., Schleper, C., Forterre, P., Garrett, R.A., and Prangishvili, D. (2009). Four newly isolated fuselloviruses from extreme geothermal environments reveal unusual morphologies and a possible interviral recombination mechanism. *Environ. Microbiol*. 11 (11), 2849-2862.

Reiter, W.D., Palm, P., Henschen, A., Lottspeich, F., and Zillig, W.D. (1987a). Identification and characterization of the genes encoding three structural proteins of the *Sulfolobus* virus-like particle SSV1. *Mol. Gen. Genet*. 206, 144-153.

Reiter, W.D., Palm, P., Yeats, S., and Zillig, W. (1987b). Gene expression in archaebacteria: physical mapping of constitutive and UV-inducible transcripts from the *Sulfolobus* virus-like particle SSV1. *Mol. Gen. Genet*. 209 (2), 270-275.

Reiter, W.D., Palm, P., and Zillig, W. (1988). Analysis of transcription in the archaebacterium *Sulfolobus* indicates that archaebacterial promoters are homologous to eukaryotic pol II promoters. *Nucleic Acids Res*. 16 (1), 1-19.

Reiter, W. Das Virusartige Partikel SSV-1 von *Sulfolobus solfataricus* Isolat B12 : UV - Induktion, Reinigung und Characterizierung [Diplomarbeit]. Tübingen: Eberhard - Karls Universität Tubingen; 1985.

Reiter, W.D., Palm, P., and Yeats, S. (1989). Transfer RNA genes frequently serve as integration sites for prokaryotic elements. *Nucleic Acids Res*. 17 (5), 1907-1914.

Ren, Y., She, Q., and Huang, L. (2013). Transcriptomic analysis of the SSV2 infection of *Sulfolobus solfataricus* with and without the integrative plasmid pSSVi. *Virology*. 441 (2), 126 - 134.

Sambrook, J., and Russell, D.W. (2001). Molecular cloning: a laboratory manual, 3rd ed. *Cold Spring Harbor laboratory Press*, Cold Spring Harbor, N.Y.

Schlenker, C., Goel, A., Tripet, B.P., Menon, S., Willi, T., Dlakić, M.,

Young, M.J., Lawrence, C.M., and Copié, V. (2012). Structural studies of E73 from a hyperthermophilic archaeal virus identify the "RH3" domain, an elaborated ribbon-helix-helix motif involved in DNA recognition. *Biochemistry*. 51(13), 2899-2910.

Scleper, C., Kubo, K., and Zillig, W. (1992). The particle SSV1 from the extremely thermophilic archaeon *Sulfolobus* is a virus: demonstration of infectivity and of transfection with viral DNA. *Proc. Natl. Acad. Sci. USA*. 89 (16), 7645-7649.

Schleper, C., Jurgens, G., and Jonuscheit, M. (2005). Genomic studies of uncultivated Archaea. *Nat. Rev. Mircobiol.* 3 (6), 478 - 488.

Schreiter, E.R., and Drennan, C.L. (2007). Ribbon-helix-helix transcription factors: variations on a theme. *Nat. Rev. Microbiol*. 5(9), 710-720.

Serre, M.C., Letzelter, C., Garel, J.R., and Duguet, M. (2002). Cleavage properties of an archaeal site-specific recombinase, the SSV1 integrase. *J. Biol. Chem.* 277 (19), 16758 - 16767.

Servin-Garcidueñas, L.E., Peng, X., Garrett, R.A., and Martínez-Romero, E. (2013). Genome sequence of a novel archaeal fusellovirus assembled from the metagenome of a mexican hot spring. *Genome Announc*. 1 (2), e0016413.

Shah, S.A., Hansen, N.R., and Garrett, R.A. (2009). Distribution of CRISPR spacer matches in viruses and plasmids of crenarchaeal acidothermophiles and implications for their inhibitory mechanism. *Biochem. Soc. Trans*. 37 (Pt 1), 23-28.

She, Q., Peng, X., Zillig, W., and Garrett, R.A. (2001a). Gene capture in archaeal chromosomes. *Nature*. 409 (6819), 478.

She, Q., Singh, R.K., Confalonieri, F., Zivanovic, Y., Allard, G., Awayez, M.J., Chan-Weiher, C.C., Clausen, I.G., Curtis, B.A., De Moors, A., Erauso, G., Fletcher, C., Gordon, P.M., Heikamp-de Jong, I., Jeffries, A.C., Kozera, C.J., Medina, N., Peng, X., Thi-Ngoc, H.P., Redder, P., Schenk, M.E., Theriault, C., Tolstrup, N., Charlebois, R.L., Doolittle, W.F., Duguet, M., Gaasterland, T., Garrett, R.A., Ragan, M.A., Sensen, C.W., Van der Oost, J. (2001b). The complete genome of the crenarchaeon *Sulfolobus solfataricus* P2. *Proc. Natl. Acad. Sci. USA*. 98 (14), 7835-7840.

She, Q., Shen, B., and Chen, L. (2004). Archaeal integrases and mechanisms of gene capture. *Biochem. Soc. Trans*. 32 (Pt 2), 222-226.

Snyder, J.C., Spulher, J., Wiedenheft, B., Roberto, F.F., Douglas, T., and Young M.J. (2004). Effects of culturing on the population structure of a hyperthermophilic virus. *Microbial. Ecol.* 48 (4), 561-566.

Snyder, J.C., Bolduc, B., and Young, M.J. (2015). 40 Years of archaeal virology: Expanding viral diversity. *Virology*. doi: 10.1016/j.virol.2015.03.031.

Sorek, R., Lawrence, C.M., and Wiedenheft, B. (2013). CRISPR-mediated adaptive immune systems in bacteria and Archaea. *Annu. Rev. Biochem*. 82, 237-266.

Stedman, K.M., Schleper, C., Rumpf, E., and Zillig, W. (1999). Genetic requirements for the function of the Archaea virus SSV1 in *Sulfolobus solfataricus*: construction and testing of viral shuttle vectors. *Genetics.* 152 (4), 1397-1405.

Stedman, K.M., She, Q., Phan, H., Arnold, H.P., Holz, I., Garrett, R.A., and Zillig W. (2003). Relationships between fuselloviruses infecting the extremely thermophilic archaeon *Sulfolobus*: SSV1 and SSV2. *Res. Microbiol.* 154 (4), 295-302.

K. M. Stedman, A. Clore and Y. Combet-Blanc. (2006). *Biogeographical Diversity of Archaeal Viruses,* In: *SGM symposium 66: Prokaryotic diversity: mechanisms and significance* (N. A. Logan, H.M. Pappin-Scott, and P.C.F. Oynston Eds). 131-144. Cambridge University Press, Cambridge.

Stedman, K.M., DeYoung, M., Saha, M., Sherman, M.B., and Morais, M. (2015). Structural insights into the architecture of the hyperthermophilic

fusellovirus SSV1. *Virology*. 474, 105-109.

Stetter, K.O. (2006). History of discovery of the first hyperthermophiles. *Extremophiles*. 10 (5), 357-362.

Sun, C.L., Barrangou, R., Thomas, B.C., Horvath, P., Fremaux, C., and Banfield J.F. (2013). Phage mutations in response to CRISPR diversification in a bacterial population. *Environ. Microbiol.* 15 (2), 463 - 470.

Veesler, D., Ng, T., Sendamarai, A., Eilers, B., Lawrence, C.M., Lok, S., Young, M.J., Jonson, J.E. and Fu, C. (2013). Atomic Structure of the 75 MDa extremophile *Sulfolobus* turreted icosahedral virus determined by CryoEM and X-ray crystallography. *Proc. Natl. Acad. Sci. USA*. 110 (14), 5504-5509.

Wang, H., Peng, N., Shah, S.A., Huang, L., and She, Q. (2015). Archaeal extrachromosomal genetic elements. *Microbial. Mol. Biol Rev.* 79 (1), 117-152.

Wang, Y., Duan, Z., Zhu, H., Guo, X., Wang, Z., Zhou, J., She, Q., and Huang, L. (2007). A novel *Sulfolobus* non-conjugative extrachromosomal genetic element capable of integration into the host genome and spreading in the presence of a fusellovirus. *Virology*. 363 (1), 124-133.

Wiedenheft, B., Stedman, K., Roberto, F., Willits, D., Gleske, A.K., Zoeller, L., Snyder, J., Douglas, T., and Young, M.J. (2004). Comparative

genomic analysis of hyperthermophilic archaeal *Fuselloviridae* viruses. *J. Virol*. 78 (4), 1954-1961.

Whitaker RJ, Grogan DW, Taylor JW. (2003). Geographic barriers isolate endemic populations of hyperthermophilic Archaea. *Science*. 2003;301:976–978.

Wirth, J.F., Snyder, J.C., Hochstein, R.A., Ortmann, A.C., Willits, D.A., Douglas, T., and Young, M.J. (2011). Development of a genetic system for the Archaea virus *Sulfolobus* turreted icosahedral virus (STIV). *Virology*. 415 (1), 6-11.

Woese C.R., and Fox, G.E. (1977). Phylogenetic structure of the prokaryotic domain: the primary kingdoms. *Proc. Natl. Acad. Sc*i. 74 (11), 5088-5090.

Xuan, J., and Feng, Y. (2012). The archaeal Sac10b protein family: conserved proteins with divergent functions. *Curr. Protein. Sci*. 13 (3), 258-266.

Xiang, X., Lanming, C., Huang, X., Luo, Y., She, Q., and Huang, L. (2005). *Sulfolobus* tengchongensis spindle-shaped virus STSV1: virus-host interactions and genomic features. *J. Virol*. 79 (14), 8677 - 8686.

Yeats, S., McWilliam, P., and Zillig, W. (1982). A plasmid in the archaeabacterium *Sulfolobus* acidocaldarius. *EMBO J*. 1 (9), 1035-1038.

Yin, Y., and Fischer, D. (2008). Identification and investigation of ORFans in the viral world. *BMC Genomics*. 9:24

Zhang, C., Cooper, T.E., Krause, D.J., and Whitaker, R.J. (2013). Augmenting the genetic toolbox for *Sulfolobus* islandicus with a stringent positive selectable marker for agmatine prototrophy. *Appl. Environ. Microbiol*. 79 (18), 5539-5549.

Zhang, J., and White, M.F. (2013). Hot and crispy: CRISPR-Cas systems in the hyperthermophile *Sulfolobus solfataricus*. *Biochem. Soc. Trans.* 41 (6), 1422-1426.

Zillig, W. (1991). Comparative biochemistry of Archaea and bacteria. *Curr Opin. Genet. Dev*. 1 (4) 544-551.

Appendix A: Transposon mutagenesis of *Sulfolobus* spindle-shaped virus 9 (SSV9)

Abstract

The vast majority of fusellovirus research has focused on *Sulfolobus* spindle-shaped virus 1 (SSV1) and has largely neglected the other viruses in this family. To expand the knowledge of the genetics from other viruses within the *Fuselloviridae*, we have initiated transposon mutagenesis studies on *Sulfolobus* spindle-shaped virus 9 (SSV9) in parallel with our work on SSV1. We show here that although SSV9 does tolerate insertion of the EZTn5 transposon (Epicentre), only one infectious insertion mutant could be obtained. For unknown reasons, less than 50% of transfections with this mutant resulted in productive infections.

Materials and Methods

Purification of SSV9 DNA from Sulfolobus

Cultures of *Sulfolobus solfataricus* strain S171 (original SSV9 host, a.k.a. G.V. #6; Wiedenheft et al., 2004) could not be reliably grown from frozen stock as could be done for SSV1 infected cells, so an alternative protocol was implemented. SSV9 DNA purified from the original SSV9 host (S171) was transformed into *S. solfataricus* S441 cells as done previously (see Chapter 3). Supernatant from this infected culture was added to approximately 50 mL of S441 cells an OD_{600nm} ~0.40-0.50 and was grown for 72 hours at 70º C with shaking. Final OD_{600nm} readings were typically ~0.7-1.0 after a 72-hour

incubation. 20 mL of this culture was spun at 3,000 x g for 10 min and the pellet was resuspended in 400 μL of resuspension buffer. SSV9 DNA was then purified following the same protocol used for SSV1 (see Chapter 3). Purified SSV9 DNA was analyzed by UV spectroscopy and EcoRI endonuclease digestion according to manufacturer's protocol (Thermo-Fisher). The remaining SSV9-infected S441 cells were spun down, resuspended in 15% glycerol and stored at -80ºC. Subsequent SSV9 DNA preps were performed in a similar fashion, however a small ice chip (ca. 50 μL) of SSV9-infected cells was added to uninfected culture instead of SSV9 supernatant.

SSV9 transposon mutagenesis reaction

The EZ-Tn5^{TM} <R6Kγori/KAN-2> Insertion Kit (Epicentre) was used to perform transposon mutagenesis on purified SSV9 DNA following the same protocol used for SSV1 DNA (See Chapter 3). An identical 30:1 ratio of SSV9 DNA to EZTn5 transposon DNA was used, as previously. The SSV9::EZTn5 reaction was transformed into pir^{+} electrocompetent *E. coli* (Epicentre) as done previously (See Chapter 3).

Transformation of Sulfolobus *and screening for virus production*

Sulfolobus strain S441 was transformed with purified EAI377 DNA as detailed elsewhere (See chapter 3). Transformed cultures were assayed for the production of infectious virus via halo assay (See Chapter 3). Viral DNA was purified from cultures that resulted in the production of a halo when spotted on

lawns of uninfected S441 using the GeneJet Plasmid Purification Kit (Thermo-Fisher). This DNA was screened for EZTn5 DNA by PCR using primers that flanked the insertion site in the SSV9 genome.

Results

Transposon mutagenesis of SSV9

Transposon mutagenesis of SSV9 was performed analogously to SSV1. The efficiency of the transposon insertion reaction was extremely high, with approximately 80% of *E. coli* transformants harboring full length SSV9::Tn5 constructs. A number of SSV9::EZTn5 mutants with insertions dispersed throughout the SSV9 genome were screened for infectivity in *Sulfolobus* S441, a strain known to be susceptible to SSV9 infection (Figure A-1 and Table A-1).

Surprisingly EAI377 with an insertion in ORF *e71*, was the only mutant found to be infectious, while all other mutants were apparently not infectious after three independent transformations. The C-terminus of ORF *e81* (EAI377) is disrupted by insertion of the transposon. Oddly, EAI377 does not reliably produce infectious virus upon transformation as indicated by the halo assay. This was true even when using the same DNA preparation that was successful in previous experiments (i.e. EAI377 DNA that had previously resulted in an infection). Furthermore, halos produced by EAI377 following transformation were much smaller than halos produced by wild type SSV9 and SSV1 (Figure A-2A).

EAI377 virions display elongated virions that are characteristic of SSV9, but not SSV1 (Figure A-2B).

Table A-1: List of SSV9 shuttle vectors in this work

Vector	Description	Reference
SSV9	Wild type SSV9, aka SSVK1	Wiedenheft et al., 2004
EAI338	SSV9::Tn5 mutant, EZ-Tn5 inserted at bp 16,924 (*vp1*)	This Work
EAI339	SSV9::Tn5 mutant, EZ-Tn5 inserted at bp 4,206 (*c158*)	This Work
EAI341	SSV9::Tn5 mutant, EZ-Tn5 inserted at bp 2,950 (*integrase*)	This Work
EAI343	SSV9::Tn5 mutant, EZ-Tn5 inserted at bp 5,647 (*b494*)	This Work
EAI348	SSV9::Tn5 mutant, EZ-Tn5 inserted at bp 5,411 (*b494*)	This Work
EAI349	SSV9::Tn5 mutant, EZ-Tn5 inserted at bp 7,774	This Work
EAI351	SSV9::Tn5 mutant, EZ-Tn5 inserted at bp 6,530 (*a460*)	This Work
EAI357	SSV9::Tn5 mutant, EZ-Tn5 inserted at bp 3,073 (*integrase*)	This Work
EAI359	SSV9::Tn5 mutant, EZ-Tn5 inserted at bp 4,979 (*b494*)	This Work
EAI368	SSV9::Tn5 mutant, EZ-Tn5 inserted at bp 4,475 (*c158*)	This Work
EAI372	SSV9::Tn5 mutant, EZ-Tn5 inserted at bp 11,955	This Work
EAI377	SSV9::Tn5 mutant, EZ-Tn5 inserted at bp 9,064 (*e81*)	This Work
EAI382	SSV9::Tn5 mutant, EZ-Tn5 inserted at bp 2,856 (*integrase*)	This Work
EAI385	SSV9::Tn5 mutant, EZ-Tn5 inserted at bp 8,644	This Work
EAI386	SSV9::Tn5 mutant, EZ-Tn5 inserted at bp 17,302 (*vp3*)	This Work
EAI389	SSV9::Tn5 mutant, EZ-Tn5 inserted at bp 10,982 (*a204*)	This Work

Transformations using similar amounts of wild type SSV9 DNA as a control consistently yielded infectious virus, indicating that the electroporation conditions are suitable and SSV9 DNA is capable of resulting in production of infectious virus. Wild type SSV9 DNA was purified from *Sulfolobus* prior to transformation, but SSV9::Tn5 DNA was purified from *E. coli*. Thus one explanation for the higher transformation is that SSV9 DNA may be modified by *Sulfolobus* and this modification results in improved transformation. To test this, EAI377 was purified from infected *Sulfolobus* cells and used to re-transform an

uninfected culture. Four independent transformations were performed and each resulted in the production of infectious virus.

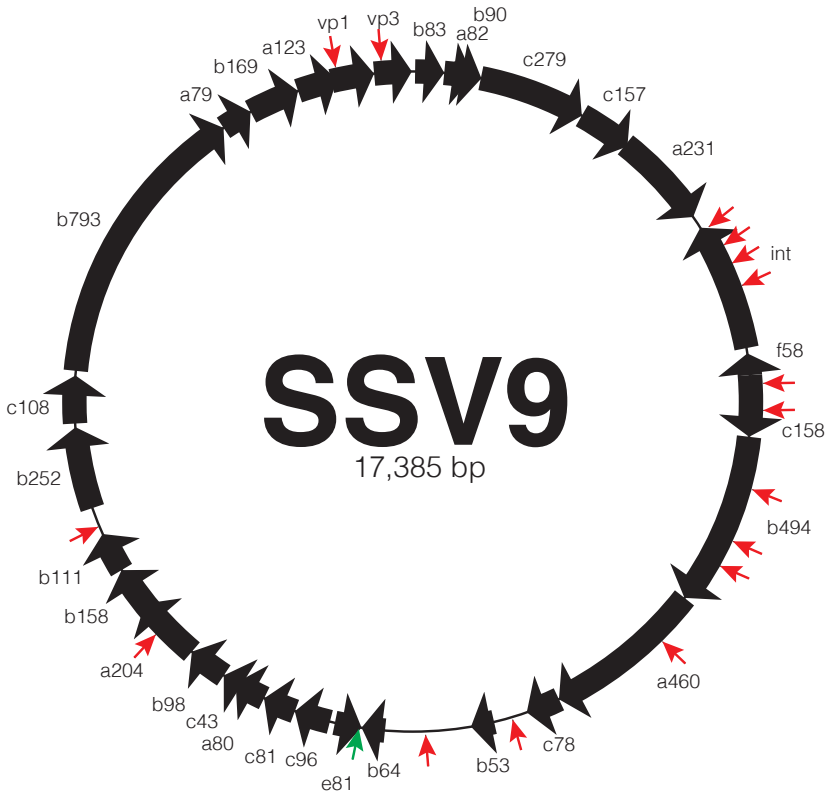


Figure A-1: Results of SSV9 mutagenesis
Map of the SSV9 genome is shown with ORFs denoted as block arrows and labeled as in Wiedenheft et al., 2004. Small arrows on the outside of the map indicate the location of EZTn5 transposon insertions in individual mutants. The color of the arrow denotes if the mutant is infectious (green) or non-infectious (red).

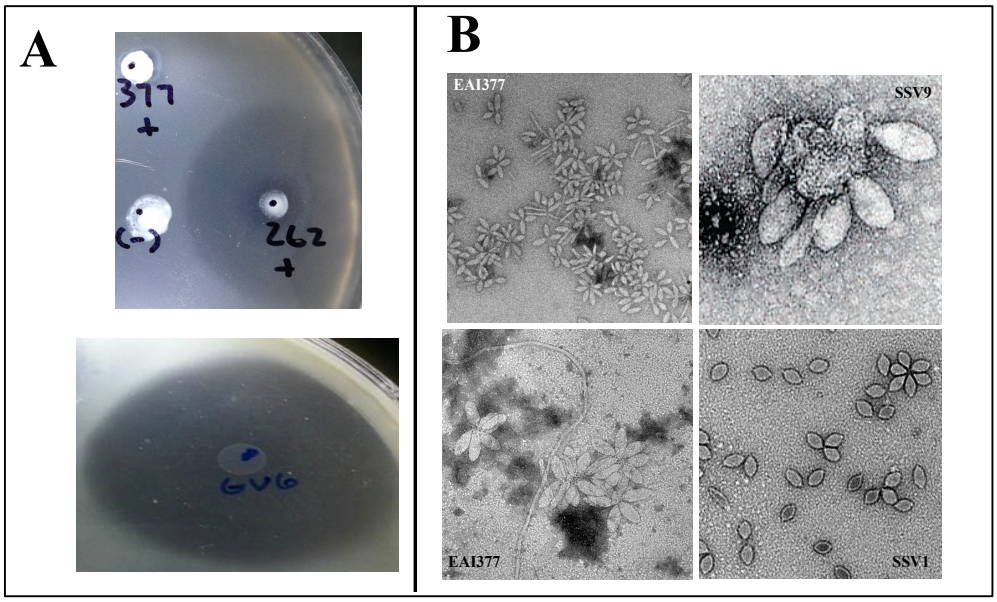


Figure A-2: Comparisons of halos produced by EAI377 and EAI377 virions
(A) Halos produced by EAI377, EAI262 (SSV1 with insertion in ORF *e96*) and SSV9 (labeled GV6 on lower plate). 5 μL of a transfected culture was spotted on a lawn of uninfected *S. solfataricus* S441 and incubated for 3 days at 75º C. An uninfected culture (-) was also spotted as a negative control.
(B) Comparison of EAI377, SSV9, and SSV1 virions. Virions were isolated from transfected cultures of *S. solfataricus* and prepared for TEM as detailed elsewhere (see Chapter 3). SSV9 image is from Wiedenheft et al., 2004.

Discussion

While SSV1 appears to be very amenable to insertional mutagenesis (Chapter 3), SSV9 appears to be almost completely intolerant. A number of mutants were isolated with insertions in a variety of ORFs ranging from universally conserved to non-conserved, however, only a single infectious mutant was able to be isolated (See Appendix B). This mutant, EAI377, has the transposon inserted at the C-terminus of ORF *e81*. E81 is poorly conserved within the *Fuselloviridae* but has a homolog in SSV1 (ORF *f93*) that has been structurally characterized and was also found to be non-essential (Kraft et al.,

2004a; Stedman et al., 1999; Figure 2). SSV1-F93 contains a wHTH motif, suggesting a role as a transcription factor in the viral life-cycle (Kraft et al., 2004a). The insertion in EAI377 is very close to the insertion of pBluescript in shuttle vector pKMSM91 that was previously shown to yield functional virus (Stedman, unpublished; Figure A-1). Apparently this region is somewhat tolerant of insertions. One of our collaborators has reported isolating an SSV9 mutant with a massive ~6 kb deletion spanning ORFs *f340* (*integrase*) to *c96*, indicating that none of these ORFs is essential for SSV9 infectivity (Whitaker, personal communication). Numerous mutants with insertions in this region were isolated yet only one of them (EAI377) was infectious. It should be noted that the host strain used by our collaborators is different and may explain these results.

It is unclear why SSV9 is so intolerant of insertional mutagenesis. Although transfection with EAI377 DNA does result in a productive infection, it was only able to do so in ~50% of transformations. Wild-type SSV9 DNA was used as a control and resulted in infection in all experiments. Furthermore, halos produced by EAI377 transformed cells were typically very small and difficult to observe (Figure A-2A). This is in stark contrast to halos produced by wild-type SSV9 which result large clearings almost entirely devoid of cells (Figure A-2A). One explanation for the reduced transformation efficiency of EAI377 compared to wild-type SSV9 could be due to packaging complications arising from the presence of an extra 2 kb of DNA. SSV9, at 17.4 kb, is already one of the largest fusellovirus genomes and may not cope well with the increased genome size

caused by insertion of the ~2 kb EZTn5. However, it is hard to imagine why this is only an issue during some transformations and not others. Another explanation is that because SSV9 DNA is purified from *Sulfolobus* prior to electroporation (whereas EAI377 DNA is purified from *E. coli*), perhaps there is a modification to the viral DNA that occurs within *Sulfolobus* that improves the efficiency of transformation. To address this, we isolated EAI377 DNA from infected *Sulfolobus* S441 cells and tested the transformation efficiency. Four independent transformations were done and each one resulted in an infection, while EAI377 purified from *E. coli* failed to produce an infection in the same experiments. EcoRI restriction endonuclease digestion of EAI377 DNA from *Sulfolobus* compared to EAI377 from *E. coli* yielded identical banding, suggesting there were not any obvious deletions or rearrangements that may have influenced the results (data not shown). Therefore, these data certainly suggest that SSV9 DNA purified from *Sulfolobus* exhibit higher transformation efficiency and this may be due to an unknown modification. Because EAI377 DNA purified from *E. coli* can successfully result in an infection, albeit at a reduced rate, it seems that this putative modification is not essential. Genome modification is not an uncommon strategy among viruses to avoid host defense systems and has been observed in the spindle-shaped virus STSV1 (Xiang et al., 2005).

Due to inconsistencies with transformation, further experimentation with SSV9 insertional mutants was postponed indefinitely.

Appendix B: Genome conservation in the *Fuselloviridae*

Introduction

Since the discovery of SSV1 in 1982 (Yeats et al., 1982), a total of 11 fuselloviruses have been isolated (see Table 1-1). A twelfth fusellovirus from a Mexican hot spring metagenome has been reported, however, a viral particle was not identified (Servin-Garcidueñas et al., 2013). All known fuselloviruses harbor genomes of similar size (14 – 17 kb) and replicate in *Sulfolobus* with the exception of Acidianus spindle-shaped virus 1 (ASV1) which infects *Acidianus* and has a significantly larger genome (~24 kb) (Redder et al., 2009).

Comparative genomics-based approaches have been useful for identifying the fusellovirus ORFs that are critical for infectivity, helping direct mutagenesis studies (Stedman et al., 2003; Wiedenheft et al., 2004; Redder et al., 2009; Held and Whitaker 2009; Chapters 2 and 3). These ORFs have been referred to as the fusellovirus core (Redder et al., 2009; Table 1-2). As more fusellovirus genomes are reported, the number of ORFs comprising the fusellovirus core shrinks. The most recent analysis using 9 fusellovirus genomes (SSV1, SSV2, SSV4-9, and ASV1) identified 13 ORFs belonging to the fusellovirus core (Redder et al., 2009). Since this publication, Stedman and colleagues have isolated two additional fuselloviruses (SSV3 and SSVL1), in addition to the previously mentioned SMF1 (Stedman et al., 2006; Servin-Garcidueñas et al.,

2013). The purpose of this appendix is to perform an updated comparative genomics analysis of the full complement of fusellovirus genomes.

Methods

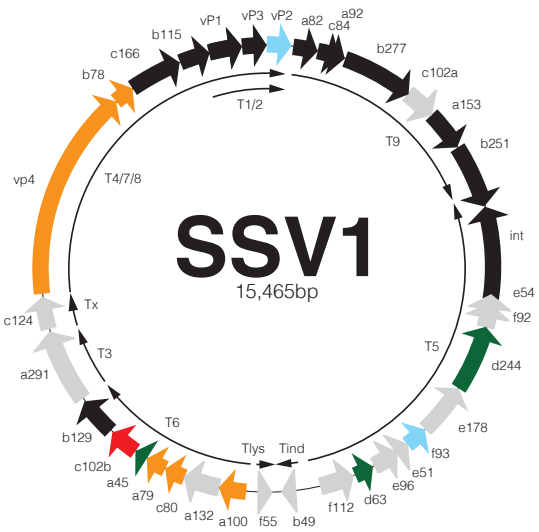
The degree of conservation for fusellovirus ORFs was determined by comparing the product of each ORF to the non-redundant database in Genbank (NCBI) using pBLAST. Homologous ORFs were identified using an arbitrary e-value cutoff of 0.01 although the vast majority of homologues had e-values which were much lower. All BLAST searches were done using default parameters. Homologous ORFs shared between SSVL and SSV3, which are not deposited in Genbank, were determined with pBLAST using the option to align two or more sequences. The recently reported SMF1 genome was analyzed but was omitted from determination of the fusellovirus core as discussed below.

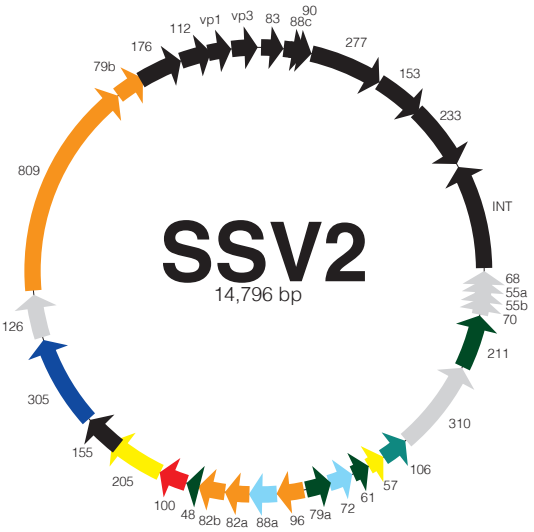
Results and discussion

As more fusellovirus genomes are discovered, the complement of ORFs comprising the fusellovirus core has gradually decreased. The fusellovirus core initially consisted of 21 ORFs following the identification of SSV2 (Stedman et al., 2003) but was recently reduced to only 13 following the identification of the ninth fusellovirus genome (Wiedenheft et al., 2004; Redder et al., 2009). We have included two additional fusellovirus genomes (SSV3 and SSVL1) in our analysis which further reduces the fusellovirus core to 12 ORFs (Table 1-1). Similar to

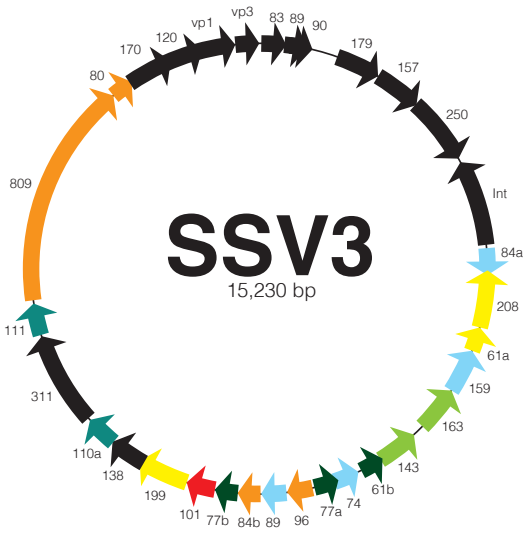
previous results, almost all of the fusellovirus core ORFs appear clustered on one half of the genome, the only exception being homologues to SSV1 ORF *b129* (Figure B-1). This new analysis has eliminated SSV1 ORF *a79* from the core, as neither of SSV3 and SSVL1 were found to encode a homolog. As in Redder et al., (2009) ORFs homologous to the putative tail fiber gene *vp4* were not judged to be part of the fusellovirus core as SSV6 and ASV1 appear to encode an alternate tail fiber module. Homologues of SSV1 ORF *c102b* are found in all but SSV6 (Figure B-1). Interestingly, 8 of the 11 fuselloviruses encode a Cas4 homolog (Figure B-1). Cas4 contains a RecB-like nuclease domain and is believed to be involved in the acquisition phase of the CRISPR-Cas pathway (Zhang et al., 2012). It is unclear why a virus would encode such a gene, but it is possible that it plays a role in DNA repair or genome replication or escape from CRISPR-Cas-mediated immunity.

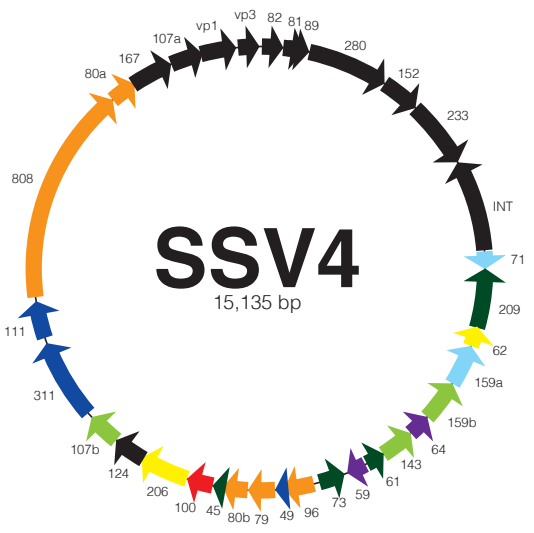
SMF1 was omitted from our analysis of the fusellovirus core for several reasons. Although the genome encodes several fusellovirus homologues, the genome is missing two core ORFs (SSV1-C84 and SSV1-B129). SMF1 also, rather peculiarly, does not encode a tail fiber module of the SSV1- or SSV6-type (Redder et al., 2009). Furthermore, other than the remaining core ORFs, SMF1 does not harbor homologs to any other fusellovirus ORF. Perhaps most crucially, a virion has not been isolated for SMF1. It is entirely possible that SMF1 is a bonafide fusellovirus but the lack of an identified virion along with several genomic abnormalities suggest otherwise.

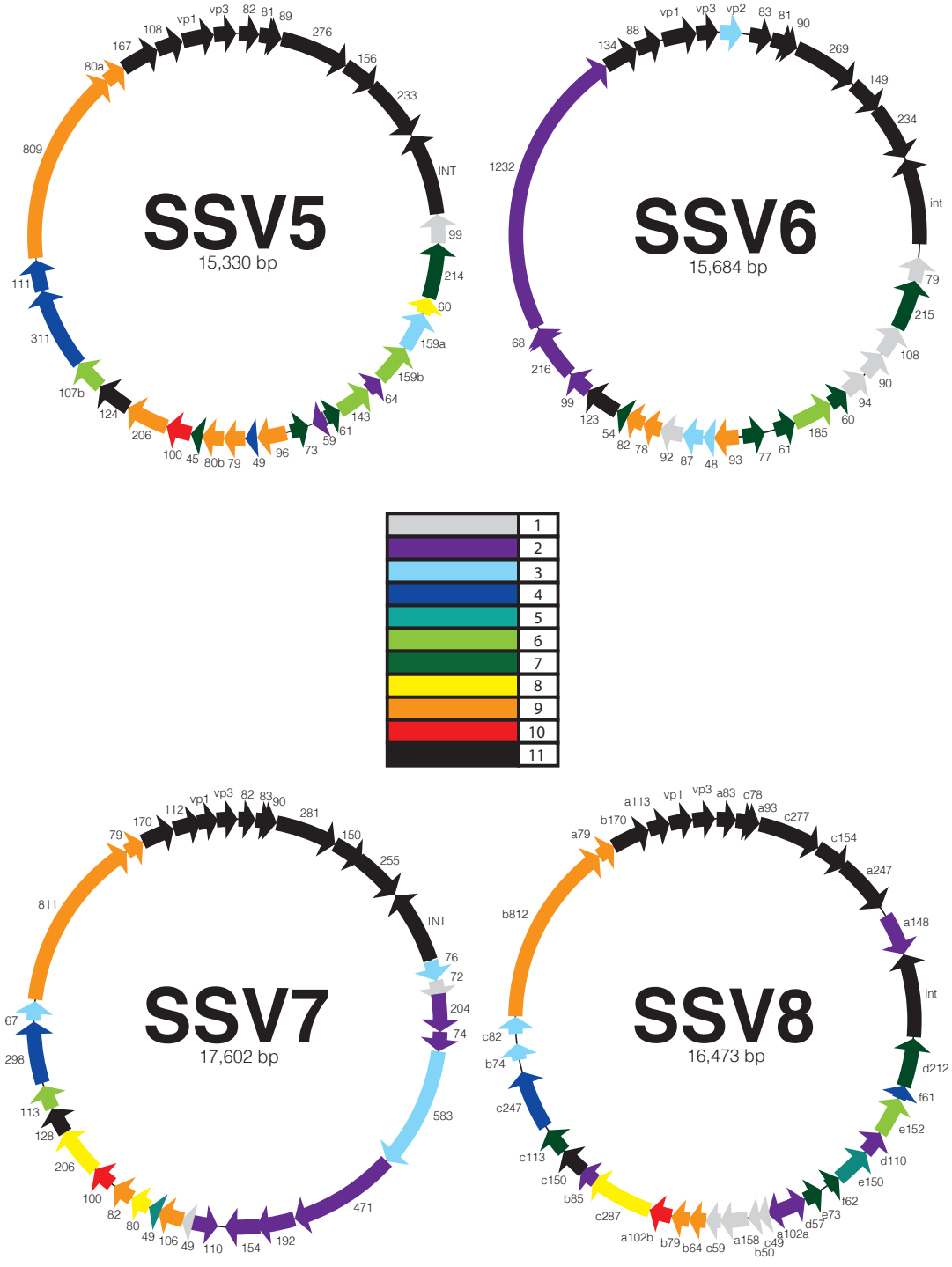

SSV1
15,465bp
vP1
vP3
vP2
a82
a92
c84
b115
c166
b78
b277
c102a
a153
b251
int
e54
f92
d244
e178
f93
e51
e96
d63
f112
b49
f55
a100
a132
c80
a79
a45
c102b
b129
a291
c124
vp4
Tx
T3
T4/7/8
T1/2
T9
T5
T6
Tlys
Tind

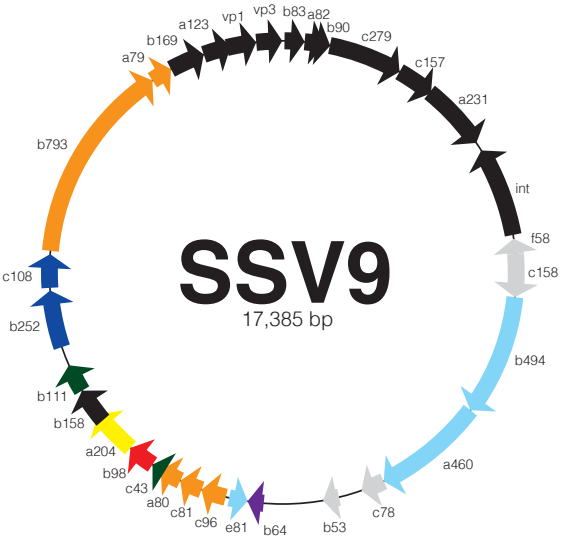

SSV2
14,796 bp
vp1
vp3
83
88c
90
112
176
79b
809
277
153
233
INT
68
55a
55b
70
211
310
106
57
61
72
79a
96
88a
82a
82b
48
100
205
155
305
126

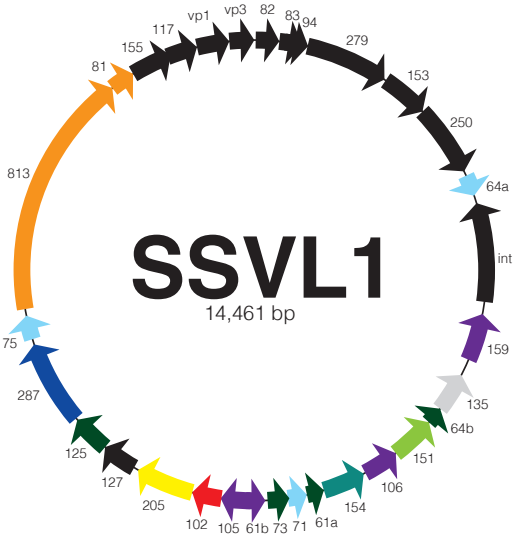

1
2
3
4
5
6
7
8
9
10
11


SSV3
15,230 bp
vp1
vp3
83
89
90
120
170
80
179
157
250
Int
84a
208
61a
159
163
143
61b
74
77a
96
89
84b
77b
101
199
138
110a
311
111
809

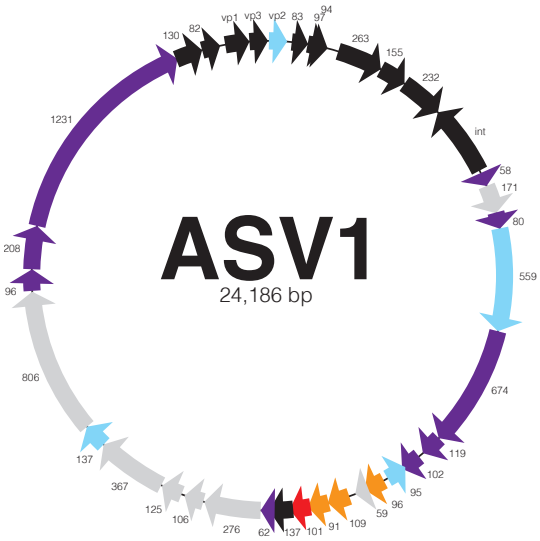

SSV4
15,135 bp
vp1
vp3
82
81
89
107a
167
80a
808
280
152
233
INT
71
209
62
159a
159b
64
143
61
59
73
96
49
79
80b
45
100
206
124
107b
311
111


SSV5
15,330 bp
SSV6
15,684 bp
SSV7
17,602 bp
SSV8
16,473 bp
1
2
3
4
5
6
7
8
9
10
11


SSV9
17,385 bp
a123
vp1
vp3
b83
a82
b90
c279
c157
a231
int
f58
c158
b494
a460
c78
b53
b64
e81
c96
c81
a80
c43
b98
a204
b158
b111
b252
c108
b793
a79
b169


SSVL1
14,461 bp
vp1
vp3
82
83
94
279
153
250
64a
int
159
135
64b
151
106
154
61a
71
73
61b
105
102
205
127
125
287
75
813
81
155
117


1
2
3
4
5
6
7
8
9
10
11


ASV1
24,186 bp
vp1
vp3
vp2
83
94
97
263
155
232
int
58
171
80
559
674
119
102
95
96
59
109
91
101
137
62
276
106
125
367
137
806
96
208
1231
130
82

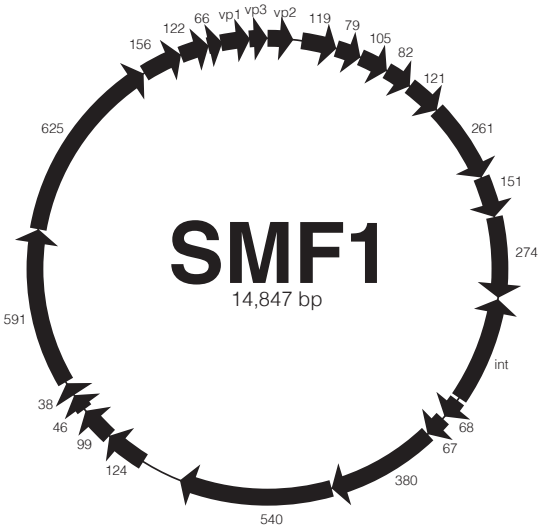

SMF1
14,847 bp
vp1
vp3
vp2
66
119
79
105
82
121
261
151
274
int
68
67
380
540
124
99
46
38
591
625
156
122

Figure B-1: Fusellovirus genome conservation
Twelve fusellovirus genomes are displayed (See Table 1-1 for references). Block arrows denote fusellovirus ORFs, colored based on the degree of conservation in the *Fuselloviridae* (see color code in figure). SSV1 transcripts are indicated and labeled (Reiter et al., 1987b; Fröls et al., 2006; Fusco et al., 2013). The genome of SMF1 is displayed but was not used in the ORF comparison.

SSV1	SSV2	SSV3	SSV4	SSV5	SSV6	SSV7	SSV8	SSV9	SSVL	ASV1	Cons
VP2					VP2					VP2	**3**
A82	83	83	82	82	83	82	A83	B83	82	83	**11**
C84	88c	89	81	81	81	83	C78	A82	83	97	**11**
A92	90	90	89	89	90	90	A83	B90	94	94	**11**
B277	277	179	280	276	269	281	C277	C279	279	263	**11**
C102a											**1**
A154	153	157	152	156	149	150	C154	C157	153	155	**11**
B251	233	250	233	233	234	255	A247	A231	250	232	**11**
INT	INT	INT	INT	INT	INT	INT	INT	INT	INT	INT	**11**
E54											**1**
F92											**1**
D244	211	208	209	214	215		D212				**7**
E178											**1**
F93								E81		95	**3**
E51											**1**
E96											**1**
D63	57	61	62	60	60		F61		64b		**7**
F112											**1**
B49											**1**
F55											**1**
A100	96	96	96	96	93	106		C96		96	**9**
A132											**1**
C80	82a	84	79	79	78	80	B64	C81			**9**
A79	82b		80	80	82	82	B79	A80		91	**9**
A45	48	77	45	45	54			C43			**7**
C102b	100	101	100	100		100	A102b	B98	102	101	**10**
B129	155	138	124	124	123	128	150	B158	185	137	**11**
A291											**1**
C124											**1**
C792	809	809	808	809		811	B812	B793	813		**9**
B78	79b	80	80a	80		79	A79	A79	81		**9**
C166	176	170	167	167	134	170	B170	B169	155	130	**11**
B115	112	120	107	108	88	112	A113	A123	117	82	**11**
VP1	VP1	VP1	VP1	VP1	VP1	VP1	VP1	VP1	VP1	VP1	**11**
VP3	VP3	VP3	VP3	VP3	VP3	VP3	VP3	VP3	VP3	VP3	**11**

Table B-1: SSV1 ORF conservation within the *Fuselloviridae*
Table view of the data presented in Figure in B-1. SSV1 ORFs are displayed in the left-most column. Homologous ORFs from the other 10 fuselloviruses are displayed in the table. Grey highlighted rows denote fusellovirus core ORFs/genes. The column on the far right (*Cons*) indicates the total number of homologues for a particular ORF.

SSV2	SSV1	SSV3	SSV4	SSV5	SSV6	SSV7	SSV8	SSV9	SSVL	ASV1	Cons
83	A82	83	82	82	83	82	A83	B83	82	83	**11**
88c	C84	89	81	81	81	83	C78	A82	83	97	**11**
90	A92	90	89	89	90	90	A83	B90	94	94	**11**
277	B277	179	280	276	269	281	C277	C279	279	263	**11**
153	A154	157	152	156	149	150	C154	C157	153	155	**11**
233	B251	250	233	233	234	255	A247	A231	250	232	**11**
Int	INT	INT	INT	INT	INT	INT	INT	INT	INT	INT	**11**
68											**1**
55a											**1**
55b											**1**
70											**1**
211	D244	208	209	214	215		D212				**7**
310											**1**
106		143	143	143			E150				**5**
57	D63	61	62	60	60		F61		64b		**8**
61		61	61	61	61		F62		61		**7**
72		74							71		**3**
79a		77	73	73	77		E73		73		**7**
96	A100	96	96	96	93	106		C96		96	**9**
88a		89			87						**3**
82a	C80	84	79	79	78	80	B64	C81			**9**
82b	A79		80	80	82	82	B79	A80		91	**9**
48	A45	77	45	45	54			C43			**7**
100	C102b	101	100	100		100	A102b	B98	102	101	**10**
205		199	206	206		206	C287	A204	205		**8**
155	B129	138	124	124	123	128	150	B158	185	137	**11**
305						298	C247		287		**4**
126											**1**
809	C792	809	808	809		811	B812	B793	813		**9**
79b	B78	80	80a	80		79	A79	A79	81		**9**
176	C166	170	167	167	134	170	B170	B169	155	130	**11**
112	B115	120	107	108	88	112	A113	A123	117	82	**11**
VP1	VP1	VP1	VP1	VP1	VP1	VP1	VP1	VP1	VP1	VP1	**11**
VP3	VP3	VP3	VP3	VP3	VP3	VP3	VP3	VP3	VP3	VP3	**11**

Table B-2: SSV2 ORF conservation within the *Fuselloviridae*
Table view of the data presented in Figure in B-1. SSV2 ORFs are displayed in the left-most column. Homologous ORFs from the other 10 fuselloviruses are displayed in the table. Grey highlighted rows denote fusellovirus core ORFs/genes. The column on the far right (*Cons*) indicates the total number of homologues for a particular ORF.

SSV3	SSV1	SSV2	SSV4	SSV5	SSV6	SSV7	SSV8	SSV9	SSVL	ASV1	Cons
83	A82	83	82	82	83	82	A83	B83	82	83	**11**
89	C84	88	81	81	81	83	C78	A82	83	97	**11**
90	A92	90	89	89	90	90	A83	B90	94	94	**11**
179	B277	277	280	276	269	281	C277	C279	279	263	**11**
157	A154	153	152	156	149	150	C154	C157	153	155	**11**
250	B251	233	233	233	234	255	A247	A231	250	232	**11**
INT	INT	INT	INT	INT	INT	INT	INT	INT	INT	INT	**11**
84a			71			76					**3**
208	D244	211	209	214	215		D212		159		**8**
61a	D63	57	62	60	60		F61		64b		**8**
159			159a	159a							**3**
163			159b	159b	185		E152		151		**6**
143		106	143	143			E150		154		**6**
61b		61	61	61	61		F62		61a		**7**
74		72							71		**3**
77		79	73	73	77		E73		73		**7**
96	A100	96	96	96	93	106		C96		96	**9**
89		88a			87						**3**
84b	C80	82	79	79	78	80	B64	C81			**9**
77	A45	48	45	45	54			C43			**7**
101	102b	100	100	100		100	A102b	B98	102	101	**10**
199		205	206	206		206	C287	A204	205		**8**
138	B129	155	124	124	123	128	150	B158	185	137	**11**
110a						113	C113	B111	125		**5**
311			311	311				B252			**4**
111	C124		111	111				C108			**5**
809	C792	809	808	809		811	B812	B793	813		**9**
80	B78	79	80a	80		79	A79	A79	81		**9**
170	C166	176	167	167	134	170	B170	B169	155	130	**11**
120	B115	112	107	108	88	112	A113	A123	117	82	**11**
VP1	VP1	VP1	VP1	VP1	VP1	VP1	VP1	VP1	VP1	VP1	**11**
VP3	VP3	VP3	VP3	VP3	VP3	VP3	VP3	VP3	VP3	VP3	**11**

Table B-3: SSV3 ORF conservation within the *Fuselloviridae*
Table view of the data presented in Figure in B-1. SSV3 ORFs are displayed in the left-most column. Homologous ORFs from the other 10 fuselloviruses are displayed in the table. Grey highlighted rows denote fusellovirus core ORFs/genes. The column on the far right (*Cons*) indicates the total number of homologues for a particular ORF.

SSV4	SSV1	SSV2	SSV3	SSV5	SSV6	SSV7	SSV8	SSV9	SSVL	ASV1	Cons
82	A82	83	83	82	83	82	A83	B83	82	83	11
81	C84	88	89	81	81	83	C78	A82	83	97	11
89	A92	90	90	89	90	90	A93	B90	94	94	11
280	B277	277	179	276	269	281	C277	C279	279	263	11
152	A154	153	157	156	149	150	C154	C157	153	155	11
233	B251	233	250	233	234	255	A247	A231	250	232	11
INT	Int	Int	Int	Int	Int	Int	Int	Int	Int	Int	11
71			84a			76					3
209	D244	211	208	214	215		D212				7
62	D63	57	61a	60	60		F61		64b		8
159a			159	159a							3
159b			163	159b	185		E152		151		6
64				64							2
143		106	143	143			E150		154		6
61		61	61b	61	61		F62		61a		7
59				59							2
73		79a	77	73	77		E73		73		7
96	A100	96	96	96	93	106		C96		96	9
49				49	48	49					4
79	C80	82	84b	79	78	80	B64	C81			9
80b	A79	82		80	82	82	B79	A80		91/109	9
45	A45	48	77	45	54			C43			7
100	C102b	100	101	100		100	A102b	B98	102	101	10
206		205	199	206		206	C287	A204	205		8
124	B129	155	124	124	123	128	C150	B158	185	137/125	11
107b				107		110/113	C113	B111	125		6
311			311	311				B252			4
111			111	111				C108			4
808	C792	809	808	809		811	B812	B793	813		9
80a	B78	79	80a	80		79	A79	A79	81		9
167	C166	176	170	167	134	170	B170	B169	155	130	11
107a	B115	112	120	108	88	112	A113	A123	117	82	11
VP1	VP1	VP1	VP1	VP1	VP1	VP1	VP1	VP1	VP1	VP1	11
VP3	VP3	VP3	VP3	VP3	VP3	VP3	VP3	VP3	VP3	VP3	11

Table B-4: SSV4 ORF conservation within the *Fuselloviridae*
Table view of the data presented in Figure in B-1. SSV4 ORFs are displayed in the left-most column. Homologous ORFs from the other 10 fuselloviruses are displayed in the table. Grey highlighted rows denote fusellovirus core ORFs/genes. The column on the far right (*Cons*) indicates the total number of homologues for a particular ORF.

SSV5	SSV1	SSV2	SSV3	SSV4	SSV6	SSV7	SSV8	SSV9	SSVL	ASV1	Cons
82	A82	83	83	82	83	82	A83	B83	82	83	11
81	C84	88	89	81	81	83	C78	A82	83	97	11
89	A92	90	90	89	90	90	A93	B90	94	94	11
276	B277	277	179	280	269	281	C277	C279	279	263	11
156	A154	153	157	152	149	150	C154	C157	153	155	11
233	B251	233	250	233	234	255	A247	A231	250	232	11
INT	INT	INT	INT	INT	INT	INT	INT	INT	INT	INT	11
99											1
214	D244	211	208	209	215		D212				7
60	D63	57	61a	62	60		F61		64b		8
159a			159	159a							3
159b			163	159b	185		E152		151		6
64				64							2
143		106	143	143			E150		154		6
61		61	61b	61	61		F62		61a		7
59				59							2
73		79	77	73	77		E73		73		7
96	A100	96	96	96	93	106		C96		96	9
49				49	48	49					4
79	C80	82	84b	79	78	80	B64	C81			9
80b	A79	82		80	82	82	B79	A80		91	9
45	A45	48	77	45	54			C43			7
100	C102b	100	101	100		100	A102b	B98	102	101	10
206		205	199	206		206	C287	A204	205		9
124	B129	155	138	124	123	128	C150	B158	185	137	11
107b				107b		110/113	C287	B111	125		6
311			311	311				B252			4
111			111	111				C108			4
809	C792	809	809	808		811	B812	B793	813		9
80a	B78	79	80	80a		79	A79	A79	81		9
167	C166	176	170	167	134	170	B170	B169	155	130	11
108	B115	112	120	107	88	112	A113	A123	117	82	11
VP1	VP1	VP1	VP1	VP1	VP1	VP1	VP1	VP1	VP1	VP1	11
VP3	VP3	VP3	VP3	VP3	VP3	VP3	VP3	VP3	VP3	VP3	11

Table B-5: SSV5 ORF conservation within the *Fuselloviridae*
Table view of the data presented in Figure in B-1. SSV5 ORFs are displayed in the left-most column. Homologous ORFs from the other 10 fuselloviruses are displayed in the table. Grey highlighted rows denote fusellovirus core ORFs/genes. The column on the far right (*Cons*) indicates the total number of homologues for a particular ORF.

SSV6	SSV1	SSV2	SSV3	SSV4	SSV5	SSV7	SSV8	SSV9	SSVL	ASV1	Cons
VP2	VP2									VP2	**3**
83	A82	83	83	82	82	82	A83	B83	82	83	**11**
81	C84	88	89	81	81	83	C78	A82	83	97	**11**
90	A92	90	90	89	89	90	A93	B90	94	94	**11**
269	B277	277	179	280	276	281	C277	C279	279	263	**11**
149	A154	153	157	152	156	150	C154	C157	153	155	**11**
234	B251	233	250	233	233	255	A247	A231	250	232	**11**
Int	INT	INT	INT	INT	INT	INT	INT	INT	INT	INT	**11**
79											**1**
215	D244	211	208	209	214		D212				**7**
108											**1**
90											**1**
94											**1**
60	D63	57	61a	62	60				64b		**7**
185			163	159	159		E152		151		**6**
61		61	61b	61	61		F62		61a		**7**
77		79a	77	73	73		E73		73		**7**
93	A100	96	96	96	96	106		C96		96	**9**
48				49		49					**3**
87		88a	89								**3**
92											**1**
78	C80	82	84b	79	79	80	B64	C81			**9**
82	A79	82		80	80	82	B79	A80		91/109	**9**
54	A45	48	77	45	45			C43			**7**
123	B129	123	138	124	124	128	C150	B158	185	137/125	**11**
99										96	**2**
216										208	**2**
68										58	**2**
1232										1231	**2**
134	C166	176	170	167	167	170	B170	B169	155	130	**11**
88	B115	112	120	107	108	112	A113	A123	117	82	**11**
VP1	VP1	VP1	VP1	VP1	VP1	VP1	VP1	VP1	VP1	VP1	**11**
VP3	VP3	VP3	VP3	VP3	VP3	VP3	VP3	VP3	VP3	VP3	**11**

Table B-6: SSV6 ORF conservation within the *Fuselloviridae*
Table view of the data presented in Figure in B-1. SSV6 ORFs are displayed in the left-most column. Homologous ORFs from the other 10 fuselloviruses are displayed in the table. Grey highlighted rows denote fusellovirus core ORFs/genes. The column on the far right (*Cons*) indicates the total number of homologues for a particular ORF.

SSV7	SSV1	SSV2	SSV3	SSV4	SSV5	SSV6	SSV8	SSV9	SSVL	ASV1	Cons
82	A82	83	83	82	82	83	A83	B83	82	83	**11**
83	C84	88	89	81	81	81	C78	A82	83	97	**11**
90	A92	90	90	89	89	90	A93	B90	94	94	**11**
281	B277	277	179	280	276	269	C277	C279	279	263	**11**
150	A154	153	157	152	156	149	C154	C157	153	155	**11**
255	B251	233	250	233	233	255	A247	A231	250	232	**11**
Int	INT	INT	INT	INT	INT	INT	INT	INT	INT	INT	**11**
76			84a	71							**3**
72											**1**
204										171	**2**
74										80	**2**
583								B494		559	**3**
471										674	**2**
192										674	**2**
154										119	**2**
110									125		**2**
49											**1**
106	A100	96	96	96	96	93		C96		96	**9**
49			84b	49	49	48					**5**
80	C80	82		79	79	78	B64	C81			**8**
82	A79	82		80	80	82	B79	A80		91/109	**9**
100	C102b	100	101	100	100		A102b	B98	102	101	**10**
206		205	199	206	206		C287	A204	205		**8**
128	B129	155	138	124	124	123	C150	B158	185	137/125	**11**
113			110a	107b	107b		C113	B111			**6**
298		305					C247		287		**4**
67							B74/C82		75		**3**
811	C792	809	809	808	809		B812	B793	813		**9**
79	B78	79	80	80	80		A79	A79	81		**9**
170	C166	176	170	167	167	134	B170	B169	155	130	**11**
112	B115	112	120	107b	108	88	A113	A123	117	82	**11**
VP1	VP1	VP1	VP1	VP1	VP1	VP1	VP1	VP1	VP1	VP1	**11**
VP3	VP3	VP3	VP3	VP3	VP3	VP3	VP3	VP3	VP3	VP3	**11**

Table B-7: SSV7 ORF conservation within the *Fuselloviridae*
Table view of the data presented in Figure in B-1. SSV7 ORFs are displayed in the left-most column. Homologous ORFs from the other 10 fuselloviruses are displayed in the table. Grey highlighted rows denote fusellovirus core ORFs/genes. The column on the far right (*Cons*) indicates the total number of homologues for a particular ORF.

SSV8	SSV1	SSV2	SSV3	SSV4	SSV5	SSV6	SSV7	SSV9	SSVL	ASV1	Cons
A83	A82	83	83	82	82	83	82	B83	82	83	**11**
C78	C84	88	89	81	81	81	83	B82	83	97	**11**
A93	A92	90	90	89	89	90	90	B90	94	94	**11**
C277	B277	277	179	280	276	269	281	C279	279	263	**11**
C154	A154	153	157	152	156	149	150	C157	153	155	**11**
A247	B251	233	250	233	233	234	255	A231	250	232	**11**
A148	E178										**2**
D355	INT	INT	INT	INT	INT	INT	INT	INT	INT	INT	**11**
D212	D244	211	208	209	214	215					**7**
F61	D63		61a	62							**4**
E152			163	159b	159b	185			151		**6**
D110									106		**2**
E150			143	143	143				154		**5**
F62		61	61b	61	61	61			61a		**7**
E73		79a	77	73	73	77			73		**7**
D57									61b		**2**
A102a									105		**2**
C49											**1**
B50											**1**
A158											**1**
C59											**1**
B64	C80	82	84b	79	79	78	80	C81			**9**
B79	A79	82		80	80	82	82	A80		109/91	**9**
A102b	C102b	100	101	100	100		100	B98	102	101	**10**
C287		205	199	206	206		206	A204	205		**8**
B85										62	**2**
C150	B129	155	138	124	124	123	128	B158	185	137/125	**11**
C113			110a	107	107		113/110	B111	125		**7**
C247		305					298/89		287		**4**
B74							67		75		**3**
C82							67			137	**3**
B812	C792	809	809	808	809		811	B793	813		**9**
A79	B78	79a	80	80	80		79	A79	81		**9**
B170	C166	176	170	167	167	134	170	B169	155	130	**11**
A113b	B115	112	120	107	108	88	112	A123	117	82	**11**
VP1	VP1	VP1	VP1	VP1	VP1	VP1	VP1	VP1	VP1	VP1	**11**
VP3	VP3	VP3	VP3	VP3	VP3	VP3	VP3	VP3	VP3	VP3	**11**

Table B-8: SSV8 ORF conservation within the *Fuselloviridae*
Table view of the data presented in Figure in B-1. SSV8 ORFs are displayed in the left-most column. Homologous ORFs from the other 10 fuselloviruses are displayed in the table. Grey highlighted rows denote fusellovirus core ORFs/genes. The column on the far right (*Cons*) indicates the total number of homologues for a particular ORF.

SSV9	SSV1	SSV2	SSV3	SSV4	SSV5	SSV6	SSV7	SSV8	SSVL	ASV1	Cons
B83	A82	83	83	82	82	83	82	A83	82	83	**11**
A82	C84	88	89	81	81	81	83	C78	83	97	**11**
B90	A92	90	90	89	89	90	90	A93	94	94	**11**
C279	B277	277	179	280	276	269	281	C277	279	263	**11**
C157	A154	153	157	152	156	149	150	C154	153	155	**11**
A231	B251	233	250	233	233	234	255	A247	250	232	**11**
INT	INT	INT	INT	INT	INT	INT	INT	INT	INT	INT	**11**
F58											**1**
C158											**1**
B494							583			559	**3**
A460							471			674	**3**
C78											**1**
B53											**1**
B64										102	**2**
E81	F93									95	**3**
C96	A100	96	96	96	96	93	106			96	**9**
C81	C80	82	84b	79	79	78	80	B64			**9**
A80	A79	82		80	80	82	82	B79		91/109	**9**
C43	A45	48	77	45	45	54					**7**
B98	C102b	100	101	100	100		100	A102b	102	101	**10**
A204		205	199	206	206		206	C287	205		**8**
B158	B129	155	138	124	124	123	128	C150	185	137	**11**
B111			110a	107b	107b		113/110	C113	125		**7**
B252			311	311	311						**4**
C108			111	111	111						**4**
B793	C792	809	809	808	809		811	B812	813		**9**
A79	B78	79	80	80a	80		79	A79	81		**9**
B169	C166	176	170	167	167	134	170	B170	155	130	**11**
A123	B115	112	120	107	108	88	112	A113	117	82	**11**
VP1	VP1	VP1	VP1	VP1	VP1	VP1	VP1	VP1	VP1	VP1	**11**
VP3	VP3	VP3	VP3	VP3	VP3	VP3	VP3	VP3	VP3	VP3	**11**

Table B-9: SSV9 ORF conservation within the *Fuselloviridae*
Table view of the data presented in Figure in B-1. SSV9 ORFs are displayed in the left-most column. Homologous ORFs from the other 10 fuselloviruses are displayed in the table. Grey highlighted rows denote fusellovirus core ORFs/genes. The column on the far right (*Cons*) indicates the total number of homologues for a particular ORF.

SSVL	SSV1	SSV2	SSV3	SSV4	SSV5	SSV6	SSV7	SSV8	SSV9	ASV1	Cons
82	A82	83	83	82	82	83	82	A83	B83	83	**11**
83	C84	88c	89	81	81	81	83	C78	A82	97	**11**
94	A92	90	90	89	89	90	90	A93	B90	94	**11**
279	B277	277	179	280	276	269	281	C277	C279	263	**11**
153	A154	153	157	152	156	149	150	C154	C157	155	**11**
250	B251	233	250	233	233	234	255	A247	A231	232	**11**
64a								B85		62	**3**
INT	INT	INT	INT	INT	INT	INT	INT	INT	INT	INT	**11**
159			208								**2**
135											**1**
64b	D63	57	61a	62	60	60					**7**
151			163	159b	159b	185		E152			**6**
106								D110			**2**
154			143	143	143			E150			**5**
61a		61	61b	61	61	61		F62			**7**
71		72	74								**3**
73		79a	77	73	73	77		E73			**7**
61b								D57			**2**
105								A102a			**2**
102	C102b	100	101	100	100		100	A102b	B98	101	**10**
205		205	199	206	206		206	C287	A204		**8**
127	B129	155	138	124	124	123	128	C150	B158	137/125	**11**
125			110a	107b	107b		110	C113	B111		**7**
287		305					298	C247			**4**
75							67	B74			**3**
813	C792	809	809	808	809		811	B812	B793		**9**
81	B78	79	80	80a	80		77	A79	A79		**9**
155	C166	176	170	167	167	134	170	B170	B169	130	**11**
117	B115	112	120	107	108	88	112	A113	A123	82	**11**
VP1	VP1	VP1	VP1	VP1	VP1	VP1	VP1	VP1	VP1	VP1	**11**
VP3	VP3	VP3	VP3	VP3	VP3	VP3	VP3	VP3	VP3	VP3	**11**

Table B-10: SSVL ORF conservation within the *Fuselloviridae*
Table view of the data presented in Figure in B-1. SSVL ORFs are displayed in the left-most column. Homologous ORFs from the other 10 fuselloviruses are displayed in the table. Grey highlighted rows denote fusellovirus core ORFs/genes. The column on the far right (*Cons*) indicates the total number of homologues for a particular ORF.

ASV1	SSV1	SSV2	SSV3	SSV4	SSV5	SSV6	SSV7	SSV8	SSV9	SSVL	Cons
VP2	VP2					VP2					**3**
83	A82	83	83	82	82	83	82	A83	B83	82	**11**
97	C84	88	89	81	81	81	83	C78	A82	83	**11**
94	A92	90	90	89	89	90	90	A93	B90	94	**11**
263	B277	277	179	280	276	269	281	C277	C279	279	**11**
155	A154	153	157	152	156	149	150	C154	C157	153	**11**
232	B251	233	250	233	233	234	255	A247	A231	250	**11**
Int	INT	INT	INT	INT	INT	INT	INT	INT	INT	INT	**11**
58							72				**2**
171											**1**
80							74				**2**
559							583		B494		**3**
674							471/192				**2**
119							154				**2**
102									B64		**2**
95	F93								E81		**3**
96	A100	96	96	96	96	93	106		C96		**9**
59											**1**
109	A79	82		80	80	82	82	B79	A80		**9**
91	A79	82		80	80	82	82	B79	A80		**9**
101	C102b	100	101	100	100		100	A102b	B98	102	**10**
137	B129	155	138	124	124	123	128	C150	B158	185	**11**
62								B85			**2**
276											**1**
106											**1**
125	B129	155	138	124	124	123	128	C150	B158	185	**11**
367											**1**
137							67	C82			**3**
806											**1**
96						99					**2**
208						213					**2**
58						68					**2**
1231						1232					**2**
130	C166	176	170	167	167	134	170	B170	B169	155	**11**
82	B115	112	120	107	108	88	112	A113	A123	117	**11**
VP1	VP1	VP1	VP1	VP1	VP1	VP1	VP1	VP1	VP1	VP1	**11**
VP3	VP3	VP3	VP3	VP3	VP3	VP3	VP3	VP3	VP3	VP3	**11**

Table B-11: ASV1 ORF conservation within the *Fuselloviridae*
Table view of the data presented in Figure in B-1. ASV1 ORFs are displayed in the left-most column. Homologous ORFs from the other 10 fuselloviruses are displayed in the table. Grey highlighted rows denote fusellovirus core ORFs/genes. The column on the far right (*Cons*) indicates the total number of homologues for a particular ORF.

SMF1	SSV1	SSV2	SSV3	SSV4	SSV5	SSV6	SSV7	SSV8	SSV9	ASV1	SSVL
VP2	VP2					VP2				VP2	
119											
79											
105											
82	A82	83	83	82	82	83	82	A83	B83	83	82
121	A92	90	90	89	89	90	90	A93	B90	94	94
261	B277	277	179	280	276	269	281	C277	C279	263	279
151	A154	153	157	152	156	149	150	C154	C157	155	153
274	B251	233	250	233	233	234	255	A247	A231	232	250
Int	INT	INT	INT	INT	INT	INT	INT	INT	INT	INT	INT
68											
67											
380											
540											
124											
99											
46											
38											
591	C792	809	809	808	809		811	B812	B793		813
625	B78	79	80	80a	80		77	A79	A79		81
156	C166	176	170	167	167	134	170	B170	B169	130	155
122	B115	112	120	107	108	88	112	A113	A123	82	117
66											
VP1	VP1	VP1	VP1	VP1	VP1	VP1	VP1	VP1	VP1	VP1	VP1
VP3	VP3	VP3	VP3	VP3	VP3	VP3	VP3	VP3	VP3	VP3	VP3

Table B-12: SMF1 homologues within the *Fuselloviridae*
SMF1 ORFs are displayed in the left-most column. Homologous ORFs from the other 11 fuselloviruses are displayed in the table.